

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 20:50:16 ; Search time 2114 Seconds
(without alignments)
6991.153 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gagtttggcaccatgtgtg.....cccgcttgaataaaaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sta.*
11: gb_sy.*
12: gb_un.*
13: gb_vi.*
14: gb_hq.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	260	100.0	AX057392 Sequence
2	260	100.0	AY243572 Synthetic
3	260	100.0	AF179612 Hepatitis
4	258.4	99.4	AR403987 Sequence
5	258.4	99.4	AX055777 Sequence
6	258.4	99.4	Y18973 Hepatitis G
7	258.4	99.4	AJ428955 Hepatitis
8	258.4	99.4	AX805211 Sequence
9	258.4	99.4	AX805212 Sequence
10	258.4	99.4	AJ277947 Hepatitis
11	257.4	99.0	AR403988 Sequence
12	257.4	99.0	AR403989 Sequence
13	257.4	99.0	AR403990 Sequence
14	257.4	99.0	AR403991 Sequence
15	257.4	99.0	AX055778 Sequence
16	257.4	99.0	AX055779 Sequence
17	257.4	99.0	AX055780 Sequence
18	257.4	99.0	AX055781 Sequence

19	82	31.5	82	6	AR404000	Sequence
20	82	31.5	82	6	AX055790	Sequence
21	34.8	13.4	229312	9	AC134603	Mus muscu
22	34.6	13.3	204413	9	AL645910	Mouse DNA
23	33.2	12.8	110000	1	AE017283	Continuation (20 o
24	33.2	12.8	182634	8	AC142353	Pan trogl
25	33.2	12.8	185516	8	AC004485	Homo sapi
26	32.8	12.6	171578	14	AC149190	Papio annu
27	32.8	12.6	180140	14	AC145521	Papio ham
28	32.8	12.6	250175	14	AC160494	Bos tauru
29	32.6	12.5	172715	9	CNS07EGP	ALS91826 BAC l3C18
30	32.6	12.5	219709	14	AC158237	Bos tauru
31	32.4	12.5	2354	13	PFM76VP1	D00056 Monkey B-ly
32	32.4	12.5	2354	13	PFM76VP1	M14494 Monkey B-ly
33	32.4	12.5	160738	14	AC161787	AC161787 Cercopithec
34	32.2	12.4	3003	6	AX780548	AX780548 Sequence
35	32.2	12.4	82414	14	AC129265	AC129265 Rattus no
36	32.2	12.4	185273	14	AC019038	AC019038 Homo sapi
37	32.2	12.4	198912	9	AC132380	AC132380 Mus muscu
38	32.2	12.4	218129	14	AC136560	AC136560 Rattus no
39	32.2	12.4	219339	14	AC121005	AC121005 Rattus no
40	32.2	12.4	234545	14	AC106571	AC106571 Rattus no
41	32.2	12.4	237041	14	AC133483	AC133483 Rattus no
42	32	12.3	3179	6	AX179748	AX179748 Sequence
43	32	12.3	67699	15	AB016815	Continuation (9 of
44	32	12.3	154622	14	AC126336	AC126336 Homo sapi
45	32	12.3	178714	14	AC129504	AC129504 Homo sapi

ALIGNMENTS

RESULT 1
AX057392
LOCUS AX057392 9399 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 1 from Patent WO0075337.
ACCESSION AX057392
VERSION AX057392.1 GI:12310132
KEYWORDS
SOURCE Hepatitis GB virus B
ORGANISM Hepatitis GB virus B
REFERENCE 1
AUTHORS Bukh, J., Yanagi, M., Emerson, S.U. and Purcell, R.H.
TITLE Infectious cdna clone of gb virus b and uses thereof
JOURNAL Patent: WO 0075337-A 1 14-DEC-2000;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES
source
1. .9399
/organism="Hepatitis GB virus B"
/mol_type="unassigned DNA"
/db_xref="taxon:39113"

ORIGIN			
Query Match 100.0%; Score 260; DB 6; Length 9399;			
Best Local Similarity 100.0%; Pred. No. 1.7e-67;			
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT	60
DB	9140	GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT	9199
QY	61	GAGTCCCTTCGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTG	120
DB	9200	GAGTCCCTTCGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTG	9259
QY	121	TGGGAAGCAGTCAGTATAATTCCTCGTGTGTGTGACGCCCTCAGACGATATTGTCCG	180
DB	9260	TGGGAAGCAGTCAGTATAATTCCTCGTGTGTGTGACGCCCTCAGACGATATTGTCCG	9319
QY	181	CTGTGTCAGACGTCAGTACCAAGGGCTGCAACCGGTTTGTTCGAAGGGAGGCAACC	240
DB	9320	CTGTGTCAGACGTCAGTACCAAGGGCTGCAACCGGTTTGTTCGAAGGGAGGCAACC	9379

/organism="Hepatitis GB virus B"
/mol_type="genomic RNA"
/specific_host="unknown"
/db_xref="taxon:39113"
/note="most closely related to hepatitis C virus;
experimental infection of tamarins causes acute resolving
hepatitis"
446..9040
/codon_start=1
/product="polyprotein"
/protein_id="AA01368.1"
/db_xref="GI:6014505"

CDS

/translation="MPVISTQSPVAPRTRKNKQTSQYSPVSIKTSVERQAKRKV
QRDAPRNYKAGTHDGLQTLAQALPAHGWGRDPRHKSRLNGLLPLGMDGT
THPLVGLVAGAVRQVQIARLEEDGNWATGFWGHLFVCLLSLACSCGARVT
DPDNTILTCORNNOVYICSPCLHPGCVICADCEWVPANPYISHPNWTGDS
FLADHIDFVMGALVTCDDALDIGELGACVGLDMLVRHLIHLIDNETGCTYLEVPTG
IDPGLFGIHWAGKVEAVIEFLTKLASQVPYALATMFSSVHYLAVALIYYASRGKY
QLLLALMIYIATSNPIRVPTGCSIAFSPCLMIPCPCHSYLSSENSEVICSPKWT
RPTILEYNNISWYPTTPGARGCMFKNNMTGCCIRNVPSTYCTMGTDVAMDTRN
TYEAGVTPMLTTHWNGSALKLAIQYPGSKEMFKPHNMGSHLYPFGSDTPIVIFY
DPVNSTLPLPERRWALPGTPPVWGSMLQVPGFYSDVKLATGLITKDKAKNKOVL
YSATGALSITGVTWKAVLILGLGSKYLILAVLCLSLCEGRASGVLPRVLPSSQ
YLOAGNDVLSQAQVAPPAFLIFPICCLRLRYAALLGFVPMAGLPLTFVFAAAQAQ
PDYDMMVLLVAGLVLMGRNHRHIALLVGPVLVALLTLHLVTPASADTEILIG
LTPPVVALVMSRFGFFFAHLLPRCALVNSYLMQRWENFWNVTLRPERFLVLVCFP
GATYDALVTFCVCHVALCLTSSAASFEGTDSRVRAHMLVRLGKCHAWYSHYLVKFF
LLVFCENGVEFYVHLHGVLPNDPASKLPLOEPPPPRFGKARVYRNBGRRLACGDTVD
GLPVVARGLDVLPAHNPDPGWAITAPTTLQCLSERGCTLSMAVWMTGIDPRWTGT
IFRUGSLATSYMGVCDNVLTYAHGSGRRRLAHPGTSIHPTITVDAANDQIYQPCG
AGSLTRCSBGTGVLVLRGLSEVNVKDDPYWCVCALPMANAKSGSAPILCSGG
HVIWFTASRNGSGSVQIRVRPLVCAQHPQYTAHATLDKTPVNEYSVQILIAPT
GSGKSLPLSWMQSKYVVLNPSVATASMPKMYHATYGVNPNCYFNKCTNTGAS
LTSYTYGVALTGACSRNDVIIICDECHATTIVLIGKVLTEAPSKNVLVLATAT
PPGVPTPHANTIELQIDBGTIPHGKIKIENELKXGHELIPEATKXCHDELNELA
RKGITAVSYRGCDLSIKIPGGDCVVVATDALCTGYTGDFDSLWVEGTCHVDLD
PTFTMGVCGVSAIVKQRRGTRGRAGIYYVDGCTPSGMVPECNIVEAFDAK
AWYGLSSTEAQOTILDYRTPGLPAIGANLDEWADLFSMVNPPEPSFVNTAKTADNYV
LLTAQLOLCHQYGAAPNDAPRMQCARLKGKCGVLRWLDGADACPGPPESEVTRYQ
MCFTEVNTSGTAAVAGVGVAMAYLAIIDFGATCYRRCSITSVPTGATVAPVUDEE
IYBECASFIPEAAVAAIDKLSTITTSPTPLETALSKLNTFLGPHATLAIIEYC
CGLVLPNPFASCVFIAGLITPLPHIKIMFLSFGAATKATDARGALVAFMAG
AAGTALGTWTSVDFVMDLGGYAAASSTACLTCLMGEMFTMDQLAGLVSAFNPA
GVVGLSACAFALTAGDPHNPRLLTMLARSNTVNEYETIATDIRKILGILEAS
TPWSISACIRHPTBDDCGLIAWGLEIWOYCNFVVICPNVLKAGVQSNVNIPOG
PYSQKQYKGPWISGMLQRCPCGABLISVENGFAKLYKGPCTCSNVRGAVPN
AKLCSARPDPDTLSLVNYGVDRYCKFERMGDHI FTVATSSPNVCTQVPTLRAA
VAVDGVQCYLGEPEKTPMTTACCYPGPGKTKVLPFRVDGHTPGVRLMQLNRDAL
ETNDCNSTNPTSDAAVSALVFKQELRNTWQLLEAISAGVDITTKLPAPSEEVVVRK
RQFRATGSLTLPPLPPSPVGPVCPSELQSPDLGPNLPPSPVLQALAMPPLLGA
GECNPPTAIGAMTETGGPDDLPSYPPKESVSEWSDSSTATTASSVVTGPPYPKI
RGKDSQAPAKPPTKKLKGSEFSCMSYTWTDVISPKTASKVLSTRAITSGLKQ
RSLVTPTEPRDLRKRQVITNRQPLFPFPPSHQVRLAKEKAKSVVWMDYDEVAH
TPSKAKSHITGLRDTGVRSGAARVAILDQKCEAGEIPSHYRTVIVPKESEVFKT
POKPTKPPRLTISPHLEMRCKVMYQGVADPVVAVMGDAYGFVDPRTRKLLSM
WSPDAGVATCDVCFDSITIPEDIMVETDIIYSAAKLSQHRAGIHTIAROLYAGGPMI
AYDREIGYRRCSRSSVYTTSSNSLITCWLKYNAAAEQAKNPRFLICGDDCTVIWK
SAGADADQAMRVFASWVKWGPAPQDCVPQPKYLELTSCSNVTSIGTSGKPYIF
LTRDPRIPLGRCSAGLGNPAAWIGILHYHPCLVWSRVLAHFMQMLFEDKLP
TVTFPMYGNVTPVEDLPSIAGHGEAFVSRYTNAELRVQSQSDTMDTPPLRA
WRKKARVLAAGAKRGGGAHAKLFLHWHATSRPLDLDKTSVARYTTTFNYCDVTSPE
GDVFTTPQRRLQKLVKVLAVIVFALGLIAVLAIS"

ORIGIN

Query Match 100.0%; Score 260; DB 13; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1.7e-67;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACATGGTCAGACCGGTTTCGGTGAGGCCATGGTCTCAAGGGGAT 60
DB 9140 GAGTTTGGCGACATGGTCAGACCGGTTTCGGTGAGGCCATGGTCTCAAGGGGAT 9199
QY 61 GACGTCCCTTCTGGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGGCTGTG 120

Db 9200 GACGTCCCTTCTGGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGGCTGTG 9259
QY 121 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACGCGCTCAGGACGTATTTGTCCG 180
Db 9260 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACGCGCTCAGGACGTATTTGTCCG 9319
QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
Db 9320 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
QY 241 CCCGCTTGGAAATTAATAACT 260
Db 9380 CCCGCTTGGAAATTAATAACT 9399
RESULT 4
AR403987 AR403987 309 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 1 from patent US 6627437.
DEFINITION AR403987
ACCESSION AR403987
VERSION AR403987.1 GI:40151915
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 309)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 1 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;
FEATURES
source 1..309
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.4%; Score 258.4; DB 6; Length 309;
Best Local Similarity 99.6%; Pred. No. 8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAGGCCATGGTCTCAAGGGGAT 60
Db 50 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAGGCCATGGTCTCAAGGGGAT 109
QY 61 GACGTCCCTTCTGGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGGCTGTG 120
Db 110 GACGTCCCTTCTGGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGGCTGTG 169
QY 121 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACGCGCTCAGGACGTATTTGTCCG 180
Db 170 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACGCGCTCAGGACGTATTTGTCCG 229
QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
Db 230 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 289
QY 241 CCCGCTTGGAAATTAATAACT 260
Db 290 CCCGCTTGGAAATTAATAACT 309
RESULT 5
AX055777
LOCUS Sequence 1 from Patent WO0073466.
DEFINITION AX055777
ACCESSION AX055777
VERSION AX055777.1 GI:12228889
KEYWORDS
SOURCE unidentified
ORGANISM unidentified

```
unclassified sequences.
1
REFERENCE
AUTHORS Traboni,C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 1 07-DEC-2000;
        ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI, S.P.A.
        (IT)
FEATURES
source Location/Qualifiers
1..309
    /organism="unidentified"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32644"
    /note="GBV-B-like virus"
1..309
    /note="cDNA complementary to the last 49 published 3'UTR
    nucleotides of GBV-B plus a novel nucleotide (C) in
    position 44 and 259 novel nucleotide sequence at the 3'
    end."
3'UTR
ORIGIN
Query Match 99.4%; Score 258.4; DB 6; Length 309;
Best Local Similarity 99.6%; Pred. No. 8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 60
Db 50 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 109
Qy 61 GACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGAAGCCATGCTCTGAAGGGGAT 120
Db 110 GACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGAAGCCATGCTCTGAAGGGGAT 169
Qy 121 TGGGAAGCAGTCAGTATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 170 TGGGAAGCAGTCAGTATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 229
Qy 181 CTGTGCAGAGCTAGTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGGGGCAACC 240
Db 230 CTGTGCAGAGCTAGTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGGGGCAACC 289
Qy 241 CCCGCTTGGAAATTAATAACT 260
Db 290 CCCGCTTGGAAATTAATAACT 309
RESULT 6
HGB18973 357 bp RNA linear VRL 19-NOV-1999
LOCUS Hepatitis GB virus B partial 3'UTR region.
DEFINITION Y18973
ACCESSION Y18973.1 GI:6018427
VERSION 3' UTR.
KEYWORDS Hepatitis GB virus B
SOURCE Hepatitis GB virus B
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE
AUTHORS Sbardellati,A., Scarselli,E., Tomei,L., Kekule,A.S. and Traboni,C.
TITLE Identification of a novel sequence at the 3' end of the GB virus B
genome
JOURNAL J. Virol. 73 (12), 10546-10550 (1999)
PUBMED 10559376
REFERENCE 2 (bases 1 to 357)
AUTHORS Traboni,C.
TITLE Direct Submission
JOURNAL Submitted (29-APR-1999) C. Traboni, IRBM P. Angeletti, Ist. di
        Ricerche di Biologia Molecolare, Via Pontina. Km. 30.600, 00040
        Pomezia, Rome, ITALY
FEATURES
source Location/Qualifiers
1..357
    /organism="Hepatitis GB virus B"
    /mol_type="genomic RNA"
    /specific host="tamarin"
    /db_xref="taxon:39113"
    .
unclassified sequences.
1
REFERENCE
AUTHORS Traboni,C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 1 07-DEC-2000;
        ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI, S.P.A.
        (IT)
FEATURES
source Location/Qualifiers
1..357
    /organism="unidentified"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32644"
    /note="GBV-B-like virus"
1..357
    /note="cDNA complementary to the last 49 published 3'UTR
    nucleotides of GBV-B plus a novel nucleotide (C) in
    position 44 and 259 novel nucleotide sequence at the 3'
    end."
3'UTR
ORIGIN
Query Match 99.4%; Score 258.4; DB 13; Length 357;
Best Local Similarity 99.6%; Pred. No. 7.8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 60
Db 98 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 157
Qy 61 GACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 120
Db 158 GACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 217
Qy 121 TGGGAAGCAGTCAGTATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 218 TGGGAAGCAGTCAGTATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 277
Qy 181 CTGTGCAGAGCTAGTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGGGGCAACC 240
Db 278 CTGTGCAGAGCTAGTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGGGGCAACC 337
Qy 241 CCCGCTTGGAAATTAATAACT 260
Db 338 CCCGCTTGGAAATTAATAACT 357
RESULT 7
GV1428955 8027 bp mRNA linear VRL 15-APR-2005
LOCUS Hepatitis GB virus B subgenomic replicon neoRepB.
DEFINITION AJ428955
ACCESSION AJ428955.1 GI:21727885
VERSION core-neo fusion protein; core-neo gene; polyprotein.
KEYWORDS Hepatitis GB virus B
SOURCE Hepatitis GB virus B
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE
AUTHORS De Tomassi,A., Pizzuti,M., Graziani,R., Sbardellati,A.,
        Altamura,S., Paonessa,G. and Traboni,C.
TITLE Cell clones selected from the Huh7 human hepatoma cell line support
        efficient replication of a subgenomic GB virus B replicon
JOURNAL J. Virol. 76 (15), 7736-7746 (2002)
PUBMED 12097587
REFERENCE 2 (bases 1 to 8027)
AUTHORS Traboni,C.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-2002) Traboni C., Biochemistry, IRBM P. Angeletti,
        Via Pontina, Km.30, 600. 00040 Pomezia (Roma), ITALY
COMMENT related sequence AJ27947.
FEATURES
source Location/Qualifiers
join(1..1281,1893..8027)
    /organism="Hepatitis GB virus B"
    /mol_type="mRNA"
    /isolate="FL3"
    /db_xref="taxon:39113"
    /focus
1282..1892
    /organism="Encephalomyocarditis virus"
    /mol_type="mRNA"
    /db_xref="taxon:12104"
1..445
    /evidence=experimental
446..1273
    /function="core-neo fusion protein"
    .
```

```
/codon_start=1
/product="neomycin phosphotransferase"
/protein_id="CAD21956.1"
/db_xref="GI:21727886"
/db_xref="GOA:O8JKE5"
/db_xref="InterPro:IPR002575"
/db_xref="UniProt/TREMBL:O8JKE5"
/translation="MPVISTOTGRAMIEQDGLHAGSPAAWVERLFYDWAQQTIGCSD
AAVFLAGQRLVFLVKDTLSGALNELQDEAARLSWLATTGVPVCAVLDVVTAEGRDW
LLGVEGQDLLSLAPAEKYSINADAMRLRLTLDPATCPFDHAKHRIIRARTRME
AGLVQDDLDHEHQGLAPAEFLPARLKARMPDGEDLVVTHGDACLINIMVENGREFSGFI
DCGRUGVADRYQDIALADIAELGEGEWADRFVLVYIGIAAPDSORIAFIARLDEFF"
1282..1892
/misc_feature
1893..7670
/note="Internal ribosome entry site (IRES)"
CDS
1893..7670
/codon_start=1
/evidence=experimental
/product="non-structural polyprotein"
/protein_id="CAD21957.1"
/db_xref="GI:21727887"
/db_xref="GOA:O8JKE4"
/db_xref="InterPro:IPR000745"
/db_xref="InterPro:IPR001410"
/db_xref="InterPro:IPR001490"
/db_xref="InterPro:IPR002166"
/db_xref="InterPro:IPR002868"
/db_xref="InterPro:IPR004109"
/db_xref="InterPro:IPR007094"
/db_xref="InterPro:IPR007095"
/db_xref="InterPro:IPR011545"
/db_xref="UniProt/TREMBL:O8JKE4"
/translation="NAPFTQLCSRLSMAVVMVTGIDPRWTGTTIPRLGLSATSY
MGFCVNDVLYTAHSGKRRLAHPITVDAANDODIYQPPCGAGSLTRCSCGE
TKGVLVSLGSLVEVNSKDDPYVCVGLPMAVAKSGSAPILCSGHVIGMTABRN
SGSGSQRVRLVCMGHPQTAHATLDTKTVNEYSVQIILAPTSGSKSTKPLS
YMOEKRYLVIIIDCECHATTATVVGKVLTEAPSKNVRLVVLATATPGVITPTPHAN
GACSNVDVIIIDCECHATTATVVGKVLTEAPSKNVRLVVLATATPGVITPTPHAN
ITETIOLTDSEGTTPFGKIKIENLKKGRHLIPEATKKHCDLDELARKIGTAVSVRG
GDIISKIPGDCVAVATDALCTGYTDPDSVDVCSLWVEGTCHVDLDPFTMGVSVRG
VSALVKGORGTGRGRAGIYYVDSCTPSGMVPECNIVEAFDAKAWYGLSSTEAQ
TILDYTRQGLPAIGANLDEWADLFWMVNPBPSFVNTAKRTADNYLTAAQLQLCH
QYGAAPADPAWQAPGKQKPGVLMRLDGDADACGPEPSEVTRIQMCFTEVNTSGI
AALAVGVGAMAYLAIDTFEGATCVRRCSITSVPGATVAVVDEEIVEECASFIPL
EMAAVADIKLSTITTPFTLETKLENTLGLGHAATILAIIEYCCGLAVLPDPNF
ASCVPFAGITTPHPKIMFSLFGGAIASKLTDARGALAFMMAAGTALGTWTS
VGFVPMPLGGYAAASTACLTFKCLMGEWPTDQLAGLYYSAFNPAGVGVLSACAM
FALTTPGDPHFNRLUTMLARNTVCNEYFIATRIDRIRKLIGILEASTPWSVISACIR
WLHTPTEDDCGLIAGLEIHWQVCPNFVVICFNLVKGAVQSMWNIIPGPFYSCOKYKG
PWIGSMLQARCPGCAELIFSVEENGFAKLYKGPRTCSNTWRGAVPNARLCSGARPD
TWTSLVNVYVGRDYCKYEKLGDIHIFVTAVSPNVCTQVPTPLRAAVADVGVQVQY
LGEPTPTWTSACCYCPDGCKTKVLPRVDGHTPGVRMLNRLDALENTDNCNTNT
PDDEAAVSALVPKQELRRNTQLLEALSAGVDTKLPAPSEVSVVRKQFARTGSLT
LPPPRSVPGVSCPSLQSDPLEGPNLPPSPVQLQAWPMLPGAGKCNCFPTAIGC
LMTETGGDDLPDLPSPYPPKVESEWSDSTATSSYVTPGPKIRKSGDQTSAPA
AKRKKLKGKLSFSCSMSTWTDVTSFKTASVLSATRAITSGLFKRLSVLVTPEPD
AELRKQKVINQPIPPPSYHKQVRLAKEKASKVGMWMDYDEVAHTPSKSAKSHIT
GLRGTDRSGAKARVLDIQQVCEAGIIPSHVROTIVIPKEEVFVTKPKTKPRL
ISYPLEMRCVEKMYGOVAVPVKAVMGDAGFDPTRFKLLSMWSPDVAVGCT
TVCFDSTIIPEDIMVETDIYSAKSDQHRAGIHTIARQLIAGGPMIAVDGREIGYR
CRSSGYVTTSSNSLTWLNKVAQAQGNRFLICGDCTVIWKSAGADADQK
RVFASWMTVMGAPQDCVPQKYSLELTSCSNVTSIGITKSGKPYFLTRDPIRGR
CSAGLGNPNSAAWGIYLIHHYPCLVWSVLAVHFEQMLFEDKLPETVTFDWMYGNV
TVPEVDLPISLIAGVHGIEAFSVVRYTNAELRVQSQSLTMDTWPPLRAWKKARVLAS
AKRGAHAKLARFLWLHATSRPLDLDKTSVARYTTFNYCDVYSEGDVFTVPQRL
QKPLVLYLAVIVFALGLIIVGLAIS"
7671..8027
/evidence=experimental
ORIGIN
Query Match 99.4%; Score 258.4; DB 13; Length 8027;
Best Local Similarity 99.6%; Pred. No. 5.2e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGTTTGGCGACCATGGTGGATCAGAACGGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
7768 GAGTTTGGCGACCATGGTGGATCAGAACGGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 7827

61 GACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
7828 GACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 7887

121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCACAGCATATTTTCTCCG 180
7888 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCACAGCATATTTTCTCCG 7947

181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
7948 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 8007

241 CCCGCTTGGAAATTAATAACT 260
8008 CCCGCTTGGAAATTAATAACT 8027

RESULT 8
AX805211
LOCUS AX805211 8069 bp RNA linear PAT 25-NOV-2003
DEFINITION Sequence 1 from Patent WO03059944.
ACCESSION AX805211
VERSION AX805211.1 GI:38522351
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1
AUTHORS de Tomassi,A., Graziani,R., Paonessa,G. and Traboni,C.
TITLE GB virus B based replicons and replicon enhanced cells
JOURNAL Patent: WO 03059944-A 1 24-JUL-2003;
INSTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.
(IT)
FEATURES
source
1..8069
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32830"
/note="GBV-B Replicon"
ORIGIN
Query Match 99.4%; Score 258.4; DB 6; Length 8069;
Best Local Similarity 99.6%; Pred. No. 5.2e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GAGTTTGGCGACCATGGTGGATCAGAACGGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
7810 GAGTTTGGCGACCATGGTGGATCAGAACGGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 7869

61 GACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
7870 GACGTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 7929

121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCACAGCATATTTTCTCCG 180
7930 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCACAGCATATTTTCTCCG 7989

181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
7990 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 8049

241 CCCGCTTGGAAATTAATAACT 260
8050 CCCGCTTGGAAATTAATAACT 8069

RESULT 9
AX805212
LOCUS AX805212 9397 bp DNA linear PAT 25-NOV-2003
```

```
DEFINITION Sequence 2 from Patent WO03059944.
ACCESSION AX805212
VERSION AX805212.1 GI:38522352
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1 de Tomassi,A., Graziani,R., Paonessa,G. and Traboni,C.
AUTHORS GB virus B based replicons and replicon enhanced cells
TITLE Patent: WO 03059944-A 2 24-JUL-2003;
JOURNAL ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.
(ITT)
FEATURES
source
1. .9397
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="GBV-B Replicon"
ORIGIN
Query Match 99.4%; Score 258.4; DB 6; Length 9397;
Best Local Similarity 99.6%; Pred. No. 5.1e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGTTTGGCAGCATGCTGGATCAGAACCGTTTCGGGTGAAGCCATGCTGAAGGGAT 60
Db 9138 GAGTTTGGCAGCATGCTGGATCAGAACCGTTTCGGGTGAAGCCATGCTGAAGGGAT 9197
QY 61 GACGTCCCTTCCTGGCTCATCCAAAACCGTCTCGGGTGGGTGAGAGTCTCGGCTGTG 120
Db 9198 GACGTCCCTTCCTGGCTCATCCAAAACCGTCTCGGGTGGGTGAGAGTCTCGGCTGTG 9257
QY 121 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTCGTCGACGCTTCACGACGATTTTGTCCG 180
Db 9258 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTCGTCGACGCTTCACGACGATTTTGTCCG 9317
QY 181 CTGTGCAGACGCTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGAGGGCAACC 240
Db 9318 CTGTGCAGACGCTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGAGGGCAACC 9377
QY 241 CCGCTTGGAAATTAATAACT 260
Db 9378 CCGCTTGGAAATTAATAACT 9397
RESULT 10
HGB277947
LOCUS HGB277947
DEFINITION Hepatitis GB virus B genomic RNA.
ACCESSION AJ277947
VERSION AJ277947.1 GI:13162187
KEYWORDS ORF1; polyprotein.
SOURCE Hepatitis GB virus B
ORGANISM Hepatitis GB virus B
REFERENCE
1 Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae.
AUTHORS Sbardellati,A., Scarselli,E., Verschoor,E., De Tomassi,A.,
Lazzaro,D. and Traboni,C.
TITLE Generation of infectious and transmissible virions from a GB virus
B full-length consensus clone in tamarins
JOURNAL J. Gen. Virol. 82 (PT 10), 2437-2448 (2001)
PUBMED 11562537
REFERENCE
2 (bases 1 to 9397)
AUTHORS Traboni,C.
TITLE Direct Submission
JOURNAL Submitted (15-MAY-2000) Traboni C., Biochemistry, IRBM P.
ANGELETTI, via Pontina km. 30.600, 00040 Pomezia (Roma), ITALY
FEATURES
source
1. .9397
Location/Qualifiers
/organism="Hepatitis GB virus B"
/mol_type="genomic RNA"
/db_xref="taxon:39113"
```

```
/clone="ACY/GBV-B/FL-3"
1..445
446..9040
/note="ORF1"
/codon_start=1
/product="polyprotein"
/protein_id="CAC33083.1"
/db_xref="GI:13162188"
/db_xref="GOA:Q99970"
/db_xref="InterPro:IPR000745"
/db_xref="InterPro:IPR001410"
/db_xref="InterPro:IPR001490"
/db_xref="InterPro:IPR002166"
/db_xref="InterPro:IPR002518"
/db_xref="InterPro:IPR002519"
/db_xref="InterPro:IPR002521"
/db_xref="InterPro:IPR002868"
/db_xref="InterPro:IPR004109"
/db_xref="InterPro:IPR007094"
/db_xref="InterPro:IPR007095"
/db_xref="InterPro:IPR011545"
/db_xref="UniProt/TrEMBL:Q99970"
/translation="MPVISTQTSVPAPRTRKNKQTOASYPVSINTSVRQGRARRKV
QRDAPRNYKIAGIHGDLQTLAALPAHGWRODPRHKSRLNGLILLDYLPLCGMDVT
THTPLVGLVAGAVVRVQVIRULEDGWNWATGFGVHLFVCLLSLACSCGARVT
DPDNTTILTNCCORNQVYICSPSTCLHEPCVICADECVPANPYIHPSNWTGDS
FLAHDIFPMGALVTCDALIGELCGACVLDGMLRWHLHIDNLTGTCYLEVTGK
IDPGLFPGMAGKVEAIFELTKLASQVPIATIMFSSVHYLVALGALITYASRQWY
QLLLALMLYIBATSNPIRVTGCSIAEFCPLMPCCHSYLSNVEVTCYSPKWT
RPVTELYNISIWYPTIPGARGCWVFKNNWTGCCRIINVPSTYCTMGTDVAMNDRN
TYEACGTPMLTTAHWNGSALKLALOYPGSKEMFKPHNMWSGHLYPEGSDTPIVVEY
DVNSTLLPPEHARLPCTPPVPGSMLOVQGFYSVDKDLATGLITDKAKNQVL
YSATGALSTGVTTRKAVVLLGLGCSKYLILAYLCYLSLFCGRASGFLRVLFSQS
YLOAGMDVLSKAQVAPFALIFICYLRLRYAALLGVFVMAAGLPTFFVAAAAAQ
PDYQWVRLVGLVWAGDRGHRILLVGFVPLVALLTLHLVTPASAFDTLVEIG
LTIPVVALVMSRFGFPAHLPRCALVNSLMORWENFWNVLPRPFRFLVLVCFP
GATYDALVFCVCHVALLTSSAASPFGTDSVRARHMLVRLGKCHAWSHYVLKFF
LLVFGENGFFYKHLHGDVLPDFASKLPLOEPFPEPEKARUVNEGERLACGDTVD
GLPVARIGDLVFAGLAMPDGNWAIPTAFITQCLSERGTSLSAMVMTGDIDPRTWGT
IFRLGSLATSYMGFVCDNVLYTAHSGKRLAHTPGTSHIPTVDAANDQIDLYPCG
AGISTRCSETGKGLVRLGSLVEVNSKDDPYMCVCGALPMVAKSGSAPQLCSG
HVGMTAARNSGKSVSQRVRLVPLVCAGYHPQYTAHATLDTKPTVNPYSQILAPT
GSKSTKLPLSYMOKYEVLPNSVATTASMPKYMATYGVNPNYVNGKCTNNGAS
LTSYTYMNLGACSRNDVLIICDECHATTATVILGIGKVLTEAPSKNVLVLATAT
PPGVIPPHANITSIQLTDEGTIPFHGKKIKENLTKGRHLIFATKXCHDELANELA
RKGTAVSYRGCDSIKPEGCVVWATDALTGTGTYDFDVSYDCLMVEGTCHVDLD
AWYGLSSEAQTILDYRTOPLGPAIGANLDEWADLFMSVNPPEPSFVNTAKRTADNY
LLTRAQLGLCHOYGYAAPNDAPRWQGBLGRKPCGVLRWLDGADACPGEPSEVTRYQ
MCFTVNTSGTAALAVGVGAMAYLAIDTFCATCVRRCHWSITSVPTGATVAPVDEE
IVECASFIPLAMVAALDKUKSTITTSPTLETALSKLMTFLGPHAAITLAIETIC
CGLVLPDNPFFASCFVAFIAGITTPPHKIKHPLSLFGGALSKLTDARGALAFWAG
AAGTALGTSVGFVDFMGLGVAASSTACLTFCMLGEMWPTMDQAGLVSAFNPA
GVVGLSACAMFALTAGDPHWPNRLLTMLARSTVCNEYFIATDRIDRKKILGLEAS
TPWSVISACIRBLHTPTDDCGLIANGLEIWOYCNPFVIFCNVLKAGVQSNVNIJPG
PFYSCQYKGPWIGSGMLQARCPGAEILISVENGFAKLYKGPRTCSNYRGAVPN
ARLCSARPDPDWTSLVYNYGRDYCKEKLGDHIFVTAVSSPFCNTPQVPTLRAA
VADVGQVQCYLGEPTKPTMTTACCYDGPDKGKTKLPFRVDGHTPGVQMNLNRDAL
ETNDCNNTNTPSPDEAAVSALVFQELRRTNQLLEAIGAGVDTTKLPASPBEVVRK
RQFRATGSLTLPPRPSVPGVSPQSLQDPLGPNLPPSPAPLAMPPLIGA
GECNPFTAIGCAMTETGGDDLPSPYPKKSVSEMSDESWTATFASVLTGPPYPKI
RGKSTQSAKAPKPTKKLKGSEFSCSNSTWTDVIFKTKASKVLVSATRAITSGFLQ
RSLVYVTEPRDAELRKKQTIINRQPLFPSPHKKVRLAKEKASKVWMDYDEVAH
TPSKSAKSHITGLDTRVSGAARAVLDLQCKEAGEIPSHYRGTIVLPKEFVVK
WSPDVGATGCTVCFDSTIPEDIMVETDIYSAAKLSQHRAGIHTIARQLVAGGPMI
KPKTKPRLISYFPLEMRCEVKMYQVAPDVVKAVMGDYGFDRVTRKLLSM
QPDVAVGATCTVCFDSTIPEDIMVETDIYSAAKLSQHRAGIHTIARQLVAGGPMI
AYDCEIGYRCRSRGSVTTSSNSLTCMLKVNAAEOAGMKNPRFLCGDDCTVIWK
WSPDVGATGCTVCFDSTIPEDIMVETDIYSAAKLSQHRAGIHTIARQLVAGGPMI
SAGADADQAMRVFASNMWKGAPQDCVPQPKYLEELTSCSNSTGTSKSKYF
LTRDPRIPLGRCSRCSGLGNFSAWIGVLIHYPCLWVSRVLAVFNMQLFEDKLPE
TVTFDYGKNTYTPVEDIPSTIAGVHGIEAFSVVRYNABEILRVSQSLTDMTPPLRA
WRKKARAVLASAKRGGAHAKLRFLLWHTSRPLDLDKTSVARYTFFNYCDVYSPE
GDVFTVTPQRLLQKFLVKYLAIVFALGIAVGLAIS"
9041..9397
3'UTR
```

ORIGIN

Query Match 99.4%; Score 258.4; DB 13; Length 9397;
 Best Local Similarity 99.6%; Pred. No. 5.1e-67;
 Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
 |||||
 Db 9138 GAGTTTGGCGACATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9197
 |||||

QY 61 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGCTGAGGAGTCTGCTGTG 120
 |||||
 Db 9198 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGCTGAGGAGTCTGCTGTG 9257
 |||||

QY 121 TGGGAAGCAGTCAGTATATATCCCGTCGTGTGTGACCGCTCAGACGTATTTGTCCG 180
 |||||
 Db 9258 TGGGAAGCAGTCAGTATATATCCCGTCGTGTGTGACCGCTCAGACGTATTTGTCCG 9317
 |||||

QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGACCCCGGTTTTTTGTTCAGCGGAGGCAACC 240
 |||||
 Db 9318 CTGTGCAGAGCGTAGTACCAAGGGCTGACCCCGGTTTTTTGTTCAGCGGAGGCAACC 9377
 |||||

QY 241 CCGCTTGGAAATTAATAACT 260
 |||||
 Db 9378 CCGCTTGGAAATTAATAACT 9397
 |||||

RESULT 11

AR403988
 LOCUS AR403988 259 bp RNA linear PAT 18-DEC-2003
 DEFINITION Sequence 2 from patent US 6627437.
 ACCESSION AR403988
 VERSION AR403988.1 GI:40151916
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 259)
 AUTHORS Traboni, C.
 TITLE GBV sequence
 JOURNAL Patent: US 6627437-A 2 30-SEP-2003;
 Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
 Pomezia;
 WO;

FEATURES

source
 1..259
 /organism="unknown"
 /mol_type="unassigned RNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
 Best Local Similarity 99.8%; Pred. No. 1.6e-66;
 Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
 |||||
 Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
 |||||

QY 62 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 121
 |||||
 Db 61 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 120
 |||||

QY 122 GGGAGCAGTCAGTATATTCCTCGTGTGTGTCAGCCCTCAGACCGTATTTGTCCGC 181
 |||||
 Db 121 GGGAGCAGTCAGTATATTCCTCGTGTGTGTCAGCCCTCAGACCGTATTTGTCCGC 180
 |||||

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACCC 241
 |||||
 Db 181 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACCC 240
 |||||

QY 242 CCGCTTGGAAATTAATAACT 260
 |||||
 Db 241 CCGCTTGGAAATTAATAACT 259
 |||||

RESULT 12

AR403989/c
 LOCUS AR403989 259 bp RNA linear PAT 18-DEC-2003
 DEFINITION Sequence 3 from patent US 6627437.
 ACCESSION AR403989
 VERSION AR403989.1 GI:40151917
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 259)
 AUTHORS Traboni, C.
 TITLE GBV sequence
 JOURNAL Patent: US 6627437-A 3 30-SEP-2003;
 Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
 Pomezia;
 WO;

FEATURES

source
 1..259
 /organism="unknown"
 /mol_type="unassigned RNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
 Best Local Similarity 99.6%; Pred. No. 1.6e-66;
 Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
 |||||
 Db 259 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 200
 |||||

QY 62 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 121
 |||||
 Db 199 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 140
 |||||

QY 122 GGGAGCAGTCAGTATATTCCTCGTGTGTGTCAGCCCTCAGACCGTATTTGTCCGC 181
 |||||
 Db 139 GGGAGCAGTCAGTATATTCCTCGTGTGTGTCAGCCCTCAGACCGTATTTGTCCGC 80
 |||||

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACCC 241
 |||||
 Db 79 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACCC 20
 |||||

QY 242 CCGCTTGGAAATTAATAACT 260
 |||||
 Db 19 CCGCTTGGAAATTAATAACT 1
 |||||

RESULT 13

AR403990
 LOCUS AR403990 259 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 4 from patent US 6627437.
 ACCESSION AR403990
 VERSION AR403990.1 GI:40151918
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 259)
 AUTHORS Traboni, C.
 TITLE GBV sequence
 JOURNAL Patent: US 6627437-A 4 30-SEP-2003;
 Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
 Pomezia;
 WO;

FEATURES

source
 1..259
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
|||||

QY 62 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 121
|||||
Db 61 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 120
|||||

QY 122 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 181
|||||
Db 121 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 180
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
|||||
Db 181 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 241 CCGCTTGGAAATTAATAACT 259
|||||

RESULT 14
AR403991/c

LOCUS AR403991 259 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 5 from patent US 6627437.
ACCESSION AR403991
VERSION AR403991.1 GI:40151919
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 259)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 5 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;

FEATURES
Location/Qualifiers
source 1..259
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 259 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 200
|||||

QY 62 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 121
|||||
Db 199 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 140
|||||

QY 122 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 181
|||||
Db 139 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 80
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
|||||
Db 79 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 19 CCGCTTGGAAATTAATAACT 1
|||||

RESULT 15
AX055778

LOCUS AX055778 259 bp RNA linear PAT 13-JAN-2001
DEFINITION Sequence 2 from Patent WO0073466.
ACCESSION AX055778
VERSION AX055778.1 GI:12228890
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 unclassified sequences.
AUTHORS Traboni, C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 2 07-DEC-2000;
ISITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI; S.P.A.
(IT)

FEATURES
Location/Qualifiers
source 1..259
/organism="unidentified"
/mol_type="unassigned RNA"
/db_xref="taxon:32644"
/note="GBV-B-like virus"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
|||||

QY 62 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 121
|||||
Db 61 AGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTGGCTGTGT 120
|||||

QY 122 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 181
|||||
Db 121 GGAAGACGTAGTATAATTCCTCGTGTGTGTGACGCCCTCAGCAGCTATTGTGCCG 180
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
|||||
Db 181 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 241 CCGCTTGGAAATTAATAACT 259
|||||

Search completed: December 7, 2005, 21:43:47
Job time : 2117 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 19:45:21 ; Search time 366 Seconds

(without alignments)
4734.481 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399

Perfect score: 260

Sequence: 1 gaggttggcgaccatgtgg.....cccgcttggaattaaaaact 260

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002bs.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	9399	4	Aaf23484
2	260	100.0	9399	12	Adj56732 Genome le
3	260	100.0	9399	12	Adj64244 GB virus
4	258.4	99.4	309	4	Aac32010 GBV-B 3'X
5	258.4	99.4	8069	9	Ada7751 Hepatitis
6	257.4	99.0	9397	9	Ada7752 Genomic h
7	257.4	99.0	259	4	Aac91998 GBV-B 3'X
8	254.4	97.8	260	12	Adj56731 3' termin
9	254.4	97.8	260	12	Adj64243 GB virus
10	248.8	95.7	362	5	Aas15946 3' nontra
11	245.4	94.4	258	4	Aac91996 GBV-B 3'X
12	108	41.5	108	12	Adj64259 GB virus
13	82	31.5	82	4	Aac92012 GBV-B 3'X
14	66	25.4	108	12	Adj56742 Mutagenic
15	66	25.4	108	12	Adj64254 Hepatitis
16	32.2	12.4	3003	10	Adf82149 Leukaemia
17	32	12.3	3179	4	Aad09560 Human tra
18	31.2	12.0	306	6	Abn25772 Human ORF
19	31.2	12.0	174155	14	Aea07466 M. ulcera

20	31	11.9	1095	6	ABS64612
21	30.6	11.8	3684	13	ADS96469
22	30.6	11.8	3772	4	ABL22533
23	30.4	11.7	522	3	AAC81741
24	30.4	11.7	2245	6	AAD33249
25	30.4	11.7	2342	3	AA93110
26	30.4	11.7	3229	8	ABZ36284
27	30.4	11.7	110000	14	ADZ12821_2
28	30.2	11.6	1012	4	AAS59679
29	30.2	11.6	1012	8	ACF64608
30	30.2	11.6	1308	4	AAS59670
31	30.2	11.6	1308	8	ACF64599
32	30	11.5	549	13	ADQ49573
33	30	11.5	2499	12	ADO07760
34	30	11.5	3855	4	ABL30105
35	30	11.5	4497	4	ABL05455
36	30	11.5	10907	4	ABL30104
37	30	11.5	13815	4	ABL05454
38	30	11.5	49999	2	AAZ23900
39	29.8	11.5	1411	13	ADX47507
40	29.8	11.5	24183	5	AAS21771
41	29.6	11.4	4659	14	ADY18412
42	29.4	11.3	460	12	ACH91975
43	29.4	11.3	1250	12	ADJ24471
44	29.4	11.3	1306	10	ADJ80204
45	29.4	11.3	1497	10	ADF09669

ALIGNMENTS

RESULT 1

AAF23484

ID AAF23484 standard; DNA; 9399 BP.

XX AAF23484;

XX AAF23484;

DT 11-SEP-2003 (revised)

DT 21-MAR-2001 (first entry)

XX GBV-B virus genome.

DE GBV-B virus genome.

XX GBV-B; hepatitis C virus; HCV; vaccine; ds.

XX Hepatitis GB virus B.

XX WO200075337-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015293.

XX 04-JUN-1999; 99US-0137694P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Bukh J, Yanagi M, Emerson SU, Purcell RH;

XX WPI; 2001-091214/10.

XX New infectious nucleic acids of the GB virus-B clone, useful for indirectly studying the molecular properties of hepatitis C virus (HCV) and in developing vaccines and therapeutics for HCV.

XX Claim 3; Page 60-63; 96pp; English.

XX The present invention relates to GB virus-B. The nucleic acid molecules of the invention are useful for indirectly studying the molecular properties of hepatitis C virus (HCV). The infectious nucleic acid sequence of the GB virus-B clone and the HCV/GBV-B chimera may be used in the development of vaccines and therapeutics for HCV. (Updated on 11-SEP-2003 to standardise OS field)

```
SQ Sequence 9399 BP; 2125 A; 2326 C; 2430 G; 2518 T; 0 U; 0 Other;

Query Match      100.0%; Score 260; DB 4; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1e-77;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAGCCATGGTCTGAAGGGAT 60
Db 9140 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAGCCATGGTCTGAAGGGAT 9199

Qy 61 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTG 120
Db 9200 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTG 9259

Qy 121 TGGGAAGCAGTCAGTATTAATCCCGTCGTGTGTGTGACGCTCTCAGACGTATTTGTCCG 180
Db 9260 TGGGAAGCAGTCAGTATTAATCCCGTCGTGTGTGTGACGCTCTCAGACGTATTTGTCCG 9319

Qy 181 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTGTCCAGCGGAGGCAACC 240
Db 9320 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTGTCCAGCGGAGGCAACC 9379

Qy 241 CCCGCTTGGAAATTAATAACT 260
Db 9380 CCCGCTTGGAAATTAATAACT 9399

RESULT 2
ADJ56732
ID ADJ56732 standard; cDNA; 9399 BP.
XX
AC ADJ56732;
XX
DT 06-MAY-2004 (first entry)
XX
DE Genome length hepatitis GB virus B cDNA SeqID 2.
XX
ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
KW vaccine; virucidal; antiinflammatory.
XX
OS Hepatitis GB virus B.
XX
FH Key Location/Qualifiers
FT stem_loop 4..21
FT /*tag= a
FT stem_loop 29..61
FT /*tag= b
FT misc_binding 63..73
FT /*tag= c
FT /*bound_moiety= "GBV-B DNA bases 236-226"
FT stem_loop 81..131
FT /*tag= d
FT stem_loop 135..154
FT /*tag= e
FT stem_loop 156..218
FT /*tag= f
FT misc_binding 226..236
FT /*tag= g
FT /*bound_moiety= "GBV-B DNA bases 73-63"
FT stem_binding 242..250
FT /*tag= h
FT /*bound_moiety= "GBV-B DNA bases 427-419"
FT misc_binding 251..257
FT /*tag= i
FT /*bound_moiety= "GBV-B DNA bases 394-388"
FT misc_binding 258..270
FT /*tag= j
FT /*bound_moiety= "GBV-B DNA bases 362-349"
FT stem_loop 273..287
FT /*tag= k
FT stem_loop 288..337
FT /*tag= l
FT stem_loop 339..348
FT ..
```

```
FT misc_binding /*tag= m
FT 349..362
FT /*tag= n
FT /*bound_moiety= "GBV-B DNA bases 270-258"
FT stem_loop 363..384
FT /*tag= o
FT misc_binding 388..394
FT /*tag= p
FT /*bound_moiety= "GBV-B DNA bases 257-251"
FT stem_loop 395..406
FT /*tag= q
FT misc_binding 410..415
FT /*tag= r
FT misc_binding 419..427
FT /*tag= s
FT /*bound_moiety= "GBV-B DNA bases 434-429"
FT stem_loop 429..434
FT /*tag= t
FT /*bound_moiety= "GBV-B DNA bases 415-410"
FT stem_loop 435..459
FT /*tag= u
XX
XX WO2004005498-A1.
XX
XX 15-JAN-2004.
XX
XX 02-JUL-2003; 2003WO-US021002.
XX
XX 03-JUL-2002; 2002US-00189359.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX (INSP ) INST PASTEUR.
XX
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX
XX WPI; 2004-091362/09.
XX
XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
XX polynucleotides. Specifically, it refers to using the hepatotropic
XX flavivirus GBV-B that has a unique phylogenetic relationship to the human
XX hepatitis C virus (HCV) and can serve as a surrogate virus in drug
XX discovery efforts related to antiviral drug development. The present
XX invention describes the construction of an infectious molecular clone
XX using the newly determined 3' terminal sequence of GBV-B. Furthermore,
XX the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
XX envelope proteins such that they can have utility as a vaccine immunogen
XX for hepatitis C. In addition, they can be used for screening compounds
XX active against viral infection, as well as for developing HCV
XX preventative and therapeutic treatments. Accordingly, these compositions
XX exhibit virucidal, antiinflammatory and hepatotropic activities. This
XX polynucleotide sequence is the GBV-B cDNA sequence of the invention.
XX
XX Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;
XX

Query Match      100.0%; Score 260; DB 12; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1e-77;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAGCCATGGTCTGAAGGGAT 60
Db 9140 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAGCCATGGTCTGAAGGGAT 9199

Qy 61 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTG 120
Db 9200 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTG 9259
```

QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 180
 DB 9260 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 9319
 QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
 DB 9320 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
 QY 241 CCGCTTGGAAATTAATAAACT 260
 DB 9380 CCGCTTGGAAATTAATAAACT 9399
 RESULT 3
 ID ADJ64244
 AC ADJ64244;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE GB virus B 3' terminal polynucleotide fragment seqid 2.
 XX
 KW antiinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
 KW hepatitis C virus; HCV; 3' terminal; ds.
 XX
 OS Hepatitis GB virus B.
 XX
 PN US2004039187-A1.
 XX
 PD 26-FEB-2004.
 XX
 PP 03-JUL-2002; 2002US-00189359.
 XX
 PR 04-JUN-1999; 99US-0137665P.
 PR 05-JUN-2000; 2000US-00587653.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (INSP) INST PASTEUR.
 XX
 PI Martin A, Sangar DV, Lemon SM, Rijnbrand R;
 XX
 DR WPI; 2004-203294/19.
 XX
 PT New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
 PT diagnosing and in treating HCV and in investigating the mechanisms for
 PT the different biological properties of the viruses.
 XX
 PS Claim 10; SEQ ID NO 2; 58pp; English.
 XX
 CC The invention describes a new isolated polynucleotide (I) encoding a 3'
 CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
 CC GBV-B genome, where at least part, but not all of a 5' nontranslated
 CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
 CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
 CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
 CC or chimeras are useful diagnosing or treating hepatitis C virus (HCV)
 CC and in investigating the mechanisms for the different biological
 CC properties of the viruses. This sequence represents a Hepatitis GB virus
 CC B (GBV-B) 3' terminal polynucleotide.
 XX
 SQ Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;
 XX
 Query Match 100.0%; Score 260; DB 12; Length 9399;
 Best Local Similarity 100.0%; Pred. No. 1e-77;
 Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGTTTGGCGACCATGTGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 60
 DB 9140 GAGTTTGGCGACCATGTGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 9199
 QY 61 GAGTCCTCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGAGTCTCTGGCTGTG 120
 DB 110 GAGTCCTCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGAGTCTCTGGCTGTG 169
 QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 180

DB 9200 GACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGAGTCTCTGGCTGTG 9259
 QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 180
 DB 9260 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 9319
 QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
 DB 9320 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
 QY 241 CCGCTTGGAAATTAATAAACT 260
 DB 9380 CCGCTTGGAAATTAATAAACT 9399
 RESULT 4
 ID AAC92010
 AC AAC92010;
 XX
 DT 20-MAR-2001 (first entry)
 XX
 DE GBV-B 3'X sequence and 3'UTR region.
 XX
 KW GBV-B; hepatitis GB viral-B; hepatitis C virus activity; ds.
 XX
 OS Hepatitis virus.
 XX
 PN WO200073466-A1.
 XX
 PD 07-DEC-2000.
 XX
 PP 22-MAY-2000; 2000WO-EP004622.
 XX
 PR 27-MAY-1999; 99GB-00012432.
 XX
 PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
 XX
 PI Traboni C;
 XX
 DR WPI; 2001-061544/07.
 XX
 PT Novel GB viral sequence useful in in vivo assays to identify agents which
 PT modulate hepatitis C virus activity.
 XX
 PS Example 2; Fig 3; 76pp; English.
 XX
 CC The present invention relates to a hepatitis GB viral-B (GBV-B)
 CC polynucleotide (see AAC91998 and AAC91996). The hepatitis GB agent was
 CC first discovered by the inoculation of tamarins with serum from a
 CC patient, whose initials were GB, affected by acute hepatitis. The serum
 CC induced hepatitis in all inoculated tamarins. The present sequence
 CC comprises the sequence in AAC91998 (the 3' terminus of GBV-B, plus the 3'
 CC UTR. This sequence confers infectivity in tamarins on otherwise non-
 CC infective GBV-B genome. This sequence is useful in in vivo assays to
 CC identify agents which modulate hepatitis C virus (HCV) activity
 XX
 SQ Sequence 309 BP; 64 A; 80 C; 100 G; 65 T; 0 U; 0 Other;
 XX
 Query Match 99.4%; Score 258.4; DB 4; Length 309;
 Best Local Similarity 99.6%; Pred. No. 1.1e-77;
 Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGTTTGGCGACCATGTGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 60
 DB 50 GAGTTTGGCGACCATGTGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 109
 QY 61 GAGTCCTCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGAGTCTCTGGCTGTG 120
 DB 110 GAGTCCTCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGAGTCTCTGGCTGTG 169
 QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGTGAGCGCTCAGACGTATTTGTCCG 180

```

Db 170 TGGGAAGCAGTCAGTATATCCCGTCGTGTGTGTGAGCCCTCAGAGGTACTTGTCCG 229
QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCAACCCCGGTTTTTGTTCGAAGCGGAGCAACC 240
Db 230 CTGTGCAGAGCGTAGTACCAAGGCTGCAACCCCGGTTTTTGTTCGAAGCGGAGCAACC 289
QY 241 CCCGCTTGAATTAATAAACT 260
Db 290 CCCGCTTGAATTAATAAACT 309

RESULT 5
ADA77751
ID ADA77751 standard; RNA; 8069 BP.
XX
AC ADA77751;
XX
DT 20-NOV-2003 (first entry)
XX
DE Hepatitis GB virus B subgenomic neo-RepD replicon RNA sequence.
XX
KW neo-RepD; GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B;
KW HCV; hepatitis C virus; antiviral; anti-HCV;
KW neomycin phosphotransferase gene; neo; ss.
XX
OS Hepatitis GB virus B.

FH Key Location/Qualifiers
FT 5'UTR 1..445
FT CDS 446..1315
FT /tag= a
FT /tag= b
FT /product= "Core-neo fusion protein"
FT misc_signal 1324..1934
FT /tag= c
FT /label= IRES
FT /note= "IRES= Internal ribosome entry site of the
FT encephalomyocarditis virus, drives translation of the GBV
FT -B NS region"
FT CDS 1935..7712
FT /tag= d
FT /product= "GBV-B polyprotein NS3-NS5B"
FT /note= "NS= non structural protein"
FT CDS 5940..7712
FT /tag= e
FT /partial
FT /product= "Non structural protein 5B (NS5B)"
FT /note= "NS5B is an RNA dependent RNA polymerase; Start
FT codon is absent"
FT 3'UTR 7710..8069
FT /tag= f

WO2003059944-A2.
PN
XX
XX
PD 24-JUL-2003.
XX
XX
PF 13-JAN-2003; 2003WO-EP000281.
XX
XX
PR 15-JAN-2002; 2002US-0348573P.
PR 06-JUN-2002; 2002US-0386655P.
XX
XX
PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
PI De Tomassi A, Graziani R, Paonessa G, Traboni C;
XX
XX
DR WPI; 2003-598503/56.
XX
XX
PT New GB virus B (GBV-B) replicon for identifying compounds that inhibit
PT GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
PT reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence,
PT and a GBV-B 3' UTR.
XX

```

```

PS Claim 1; Fig 1; 81pp; English.
XX
CC This invention relates to a novel GB virus-B (GBV-B) replicon and
CC replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
CC autonomously replicate in a cultured cell to produce detectable levels of
CC one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5',
CC UTR, GBV-B structural region, selection or reporter sequence, internal
CC ribosome entry site, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
CC they are useful in providing tools for studying GBV-B replication,
CC polyprotein production and processing, identifying compounds that inhibit
CC GBV-B, providing a surrogate model for identifying compounds that inhibit
CC HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
CC Due to the similarity between GBV-B and the hepatitis C virus (HCV),
CC compounds that inhibit GBV-B may be useful antiviral agents, specifically
CC anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
CC neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
CC replacing the regions coding for structural proteins and the NS2 protein
CC with the sequences of neomycin phosphotransferase gene (neo) and
CC encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
CC the plasmid FL3/pACYC177. This polynucleotide sequence is the subgenomic
CC neo-RepD replicon sequence of the invention.
XX
SQ Sequence 8069 BP; 1863 A; 2075 C; 2136 G; 0 T; 1995 U; 0 Other;

Query Match 99.4%; Score 258.4; DB 9; Length 8069;
Best Local Similarity 76.2%; Pred. No. 3.4e-77;
Matches 198; Conservative 61; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGTGTGATCAGAACCCGTTTCGGGTGAAGCCCATGCTCTGAAGGGAT 60
Db 7810 GAGUUGGCGACCAUGGUGGAUCAAGACCGUUGCGGUGAAGCCCAUGGUCUGAAGGGAU 7869
QY 61 GACGTCCCTTCTGGCTCATCCACAAAACCCCTCTCGGTGGTGAGGAGTCTGCTGTG 120
Db 7870 GACGUCCUUGGCGUCAUCCACAAAACCCGUCUGGUGGAGGAGGUGGUGGUGGUG 7929
QY 121 TGGGAAGCAGTCAGTATATCCCGTCGTGTGTGTGACGCTCAGCAGCTATTGTGTCG 180
Db 7930 UGGGAGAGCAGUCAGUAAUUAUCCGUGUGUGAGAGCCGUCACGACGACUAGUUGCCG 7989
QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGTTTTTGTTCGAAGCGGAGCAACC 240
Db 7990 CUGUGCAGAGCGUAGUACCAAGGCGUGCACCCCGUUUUUUGUCCAAGCGGAGGCAACC 8049
QY 241 CCCGCTTGAATTAATAAACT 260
Db 8050 CCCGCTTGAATTAATAAACT 8069

RESULT 6
ADA77752
ID ADA77752 standard; cDNA; 9397 BP.
XX
AC ADA77752;
XX
DT 20-NOV-2003 (first entry)
XX
DE Genomic hepatitis GB virus B (GBV-B) replicon cDNA sequence.
XX
KW GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B; HCV;
KW hepatitis C virus; antiviral; anti-HCV; neomycin phosphotransferase gene;
KW neo; ss.
XX
OS Hepatitis GB virus B.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..445
FT CDS 446..9040
FT /tag= a
FT /tag= b
FT /product= "GBV-B polyprotein"
FT /note= "From core protein to non structural protein 5B"
FT CDS 7268..9040

```

```

FT FT /*tag= C
FT FT /partial
FT FT /product= "Non structural protein 5B (NS5B)"
FT FT /note= "NS5B is an RNA dependent RNA polymerase; start
FT FT codon is absent"
FT FT 9038..9397
FT FT /*tag= d
XX XX WO2003059944-A2.
XX XX 24-JUL-2003.
XX XX 13-JAN-2003; 2003WO-EP000281.
XX XX 15-JAN-2002; 2002US-0348573P.
XX XX 06-JUN-2002; 2002US-0386655P.
XX XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX XX De Tomassi A, Graziani R, Paonessa G, Traboni C;
XX XX WPI: 2003-598503/56.
XX XX P-PSDB; ADA77730.
XX XX New GB virus B (GBV-B) replicon for identifying compounds that inhibit
XX XX GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
XX XX reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence,
XX XX and a GBV-B 3' UTR.
XX XX Claim 3; Fig 2; 81pp; English.
XX XX This invention relates to a novel GB virus-B (GBV-B) replicon and
XX XX replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
XX XX autonomously replicate in a cultured cell to produce detectable levels of
XX XX one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5'
XX XX UTR, GBV-B structural region, selection or reporter sequence, internal
XX XX ribosome entry site, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
XX XX they are useful in providing tools for studying GBV-B replication,
XX XX polypeptide production and processing, identifying compounds that inhibit
XX XX GBV-B, providing a surrogate model for identifying compounds that inhibit
XX XX HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
XX XX Due to the similarity between GBV-B and the hepatitis C virus (HCV),
XX XX compounds that inhibit GBV-B may be useful antiviral agents, specifically
XX XX anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
XX XX neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
XX XX replacing the regions coding for structural proteins and the NS2 protein
XX XX with the sequences of neomycin phosphotransferase gene (neo) and
XX XX encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
XX XX the plasmid FL3/pACYC177. This polynucleotide sequence is the genomic GBV
XX XX -B replicon cDNA sequence of the invention.
XX XX Sequence 9397 BP; 2125 A; 2331 C; 2431 G; 2510 T; 0 U; 0 Other;
SQ
Query Match 99.4%; Score 258.4; DB 9; Length 9397;
Best Local Similarity 99.6%; Pred. No. 3.6e-77;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9138 GAGTTTGGCGACCATGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9197
QY 61 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTG 120
DB 9198 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTG 9257
QY 121 TGGGAACAGTCAGTATAATTTCCCGTCGTGTGGTGACGCCCTCACACACGTATTTGTCCG 180
DB 9258 TGGGAACAGTCAGTATAATTTCCCGTCGTGTGGTGACGCCCTCACACACGTATTTGTCCG 9317
QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGGCAACC 240
DB 9318 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGGCAACC 9377

```

```

QY 241 CCCGCTTGGAAATTAATAAACT 260
DB 9378 CCCGCTTGGAAATTAATAAACT 9397

RESULT 7
AAC91998
ID AAC91998 standard; DNA; 259 BP.
XX AAC91998;
XX 20-MAR-2001 (first entry)
XX GBV-B 3' X DNA sequence.
XX GBV-B; hepatitis GB viral-B; hepatitis C virus activity; ds.
XX Hepatitis virus.
XX WO200073466-A1.
XX 07-DEC-2000.
XX 22-MAY-2000; 2000WO-EP004622.
XX 27-MAY-1999; 99GB-00012432.
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX Traboni C;
XX WPI: 2001-061544/07.
XX Novel GB viral sequence useful in in vivo assays to identify agents which
XX modulate hepatitis C virus activity.
XX Claim 1; Page 14; 76pp; English.
XX The present sequence is an isolated hepatitis GB viral-B (GBV-B)
XX polynucleotide. The hepatitis GB agent was first discovered by the
XX inoculation of tamarins with serum from a patient, whose initials were
XX GB, affected by acute hepatitis. The serum induced hepatitis in all
XX inoculated tamarins. The present sequence comprises the 3' terminus of
XX GBV-B, and confers infectivity in tamarins on otherwise non-infective GBV
XX -B genome. This sequence is useful in in vivo assays to identify agents
XX which modulate hepatitis C virus (HCV) activity
XX Sequence 259 BP; 53 A; 64 C; 81 G; 61 T; 0 U; 0 Other;
SQ
Query Match 99.0%; Score 257.4; DB 4; Length 259;
Best Local Similarity 99.6%; Pred. No. 2.2e-77;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
DB 1 AGTTTGGCGACCATGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
QY 62 ACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 121
DB 61 ACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 120
QY 122 GGGGAACAGTCAGTATAATTTCCCGTCGTGTGGTGACGCCCTCACACACGTATTTGTCCGC 181
DB 121 GGGGAACAGTCAGTATAATTTCCCGTCGTGTGGTGACGCCCTCACACACGTATTTGTCCGC 180
QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGGCAACC 241
DB 181 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGGCAACC 240
QY 242 CCCTTTGGAAATTAATAAACT 260
DB 241 CCCTTTGGAAATTAATAAACT 259

```

```
RESULT 8
ADJ56731
ID ADJ56731 standard; RNA; 260 BP.
XX AC
XX ADJ56731;
XX DT
XX 06-MAY-2004 (first entry)
XX DE
XX 3' terminal sequence of the hepatitis GB virus B RNA SeqID 1.
XX ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
XX KW vaccine; virucidal; antiinflammatory.
XX OS Hepatitis GB virus B.
XX PH Key Location/Qualifiers
XX FT misc_binding 152..156
XX FT /tag= a
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 199-195"
XX FT misc_binding 157..160
XX FT /tag= b
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 192-189"
XX FT misc_binding 161..162
XX FT /tag= c
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 187-186"
XX FT misc_binding 163..166
XX FT /tag= d
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 184-181"
XX FT stem_loop 168..179
XX FT /tag= e
XX FT /label= SL-3
XX FT misc_binding 181..184
XX FT /tag= f
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 166-163"
XX FT misc_binding 186..187
XX FT /tag= g
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 162-161"
XX FT misc_binding 189..192
XX FT /tag= h
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 160-157"
XX FT misc_binding 195..199
XX FT /tag= i
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 156-152"
XX FT stem_loop 201..212
XX FT /tag= j
XX FT /label= SL-2
XX FT misc_binding 214..220
XX FT /tag= k
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 258-252"
XX FT misc_binding 221..230
XX FT /tag= l
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 250-241"
XX FT stem_loop 232..239
XX FT /tag= m
XX FT misc_binding 241..250
XX FT /label= SL-1
XX FT /tag= n
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 230-221"
XX FT misc_binding 252..258
XX FT /tag= o
XX FT /bound_moiety= "3' terminal GBV-B RNA bases 220-214"
XX WO2004005498-A1.
XX PN
XX 15-JAN-2004.
XX 02-JUL-2003; 2003WO-US021002.
XX 03-JUL-2002; 2002US-00189359.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX (INSP ) INST PASTEUR.
XX PA
```

```
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX WPI; 2004-091362/09.
XX New chimeric GBV-B polynucleotide, useful as a model for hepatitis C
XX virus, for identifying compounds active against a viral infection, or for
XX developing hepatitis C virus preventive and therapeutic treatments.
XX Example 1; SEQ ID NO 1; 108pp; English.
XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
XX polynucleotides. Specifically, it refers to using the hepatotropic
XX flavivirus GBV-B that has a unique phylogenetic relationship to the human
XX hepatitis C virus (HCV) and can serve as a surrogate virus in drug
XX discovery efforts related to antiviral drug development. The present
XX invention describes the construction of an infectious molecular clone
XX using the newly determined 3' terminal sequence of GBV-B. Furthermore,
XX the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
XX envelope proteins such that they can have utility as a vaccine immunogen
XX for hepatitis C. In addition, they can be used for screening HCV
XX active against viral infection, as well as for developing HCV
XX preventative and therapeutic treatments. Accordingly, these compositions
XX exhibit virucidal, antiinflammatory and hepatotropic activities. This
XX polynucleotide sequence is the 3' terminal sequence of GBV-B RNA sequence
XX of the invention.
XX Sequence 260 BP; 53 A; 63 C; 82 G; 0 T; 62 U; 0 Other;
XX Query Match 97.8%; Score 254.4; DB 12; Length 260;
XX Best Local Similarity 75.8%; Pred. No. 2.3e-76;
XX Matches 194; Conservative 61; Mismatches 1; Indels 0; Gaps 0;
QY 5 TTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACG 64
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 4 UUGGCGCAUUGGUGGUAUCAGAACCGUUCGCGGUGAAGCCAUUGGUCUAGAAGGGGAUGCG 63
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 65 TCCCTTCCTGGCTCATCCACAAAAACCGTCTCGGGTGGTGAGGAGTCTCGGCTGTGTGGG 124
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 64 UCCCCUUGGUGGUAUCCGCAAAAAACCGUCUGGUGGUGAGAGUCCUGGUCUGUGUGGG 123
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 125 AAGCAGTCAGTATATATCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 184
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 124 AAGCAGUCAGUUAUUAUCCCGUGUGUGUGAGCGCCUACAGAGUUAUUGUCCGUGU 183
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 185 GCAGAGCGTAGTACCAAGGGTGCACCCCGTTTGTTCAGCGGAGGGCAACCCCG 244
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 184 GCAGAGCGUAGUACCAAGGCGUCCACCCCGUUUUUUGUCCAAAGCGGAGGCAACCCCG 243
Db ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 245 CTTGGAATTAATAAACT 260
Db ::::::::::::::::::::::
QY 244 CUUGGAUUUAAAAACU 259
Db ::::::::::::::::::::::
RESULT 9
ADJ64243
ID ADJ64243 standard; DNA; 260 BP.
XX AC
XX ADJ64243;
XX DT
XX 20-MAY-2004 (first entry)
XX DE
XX GB virus B 3' terminal polynucleotide fragment seqid 1.
XX antiinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
XX hepatitis C virus; HCV; 3' terminal; ds.
XX Hepatitis GB virus B.
XX OS
XX US2004039187-A1.
XX PN
XX 26-FEB-2004.
XX PD
XX
```


XX The invention describes a new isolated polynucleotide (I) encoding a 3' CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric CC GBV-B genome, where at least part, but not all of a 5' nontranslated CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR. CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides CC or chimaeras are useful diagnosing or treating hepatitis C virus (HCV) CC and in investigating the mechanisms for the different biological CC properties of the viruses. This sequence represents a Hepatitis GB virus CC B (GBV-B) 3' terminal polynucleotide. Note: This sequence represents CC nucleotides 152-259 of the sequence shown in ADJ64243.

XX Sequence 108 BP; 23 A; 30 C; 30 G; 0 T; 25 U; 0 Other;

Query Match 41.5%; Score 108; DB 12; Length 108;
Best Local Similarity 76.9%; Pred. No. 1.7e-26;
Matches 83; Conservative 25; Mismatches 0; Indels 0; Gaps 0;

QY 153 TGGTGACGCTCAGACGATTGTCGCTGTGCAGCGTAGTACCAAGGGCTGCACCC 212
DB 1 UGUGAGCGCUACGACGUUUUUGCGUGGACGAGCGUAGUACCAAGGGCGCACCC 60

QY 213 CGGTTTTGTTCCAGCGAGGCGCAACCCCGCTTGAATTAATAACT 260
DB 61 CGGUUUUGUCCAGCGAGGCGCAACCCCGCGUUGGAUUAUAAACU 108

RESULT 13

AAC92012
ID AAC92012 standard; RNA; 82 BP.

XX AAC92012;

XX 20-MAR-2001 (first entry)

XX GBV-B 3'X RNA sequence #2.

XX GBV-B; hepatitis GB viral-B; ds.

XX Hepatitis virus.

XX WO200073466-A1.

XX 07-DSC-2000.

XX 22-MAY-2000; 2000WO-EP004622.

XX 27-MAY-1999; 99GB-00012432.

XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX Traboni C;

XX WPI; 2001-061544/07.

XX Novel GB viral sequence useful in in vivo assays to identify agents which PT modulate hepatitis C virus activity.

XX Disclosure; Fig 4; 76pp; English.

XX The present sequence is an isolated hepatitis GB viral-B (GBV-B) CC polynucleotide. The hepatitis GB agent was first discovered by the CC inoculation of tamarins with serum from a patient, whose initials were CC GB, affected by acute hepatitis. The serum induced hepatitis in all CC inoculated tamarins. The present sequence comprises the 3' terminus of CC GBV-B

XX Sequence 82 BP; 19 A; 23 C; 23 G; 0 T; 17 U; 0 Other;

Query Match 31.5%; Score 82; DB 4; Length 82;
Best Local Similarity 79.3%; Pred. No. 1.1e-17;
Matches 65; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

QY 179 CGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAA 238
DB 1 CGCUGUGCAGAGCGUAGUACCAAGGGCGUGCACCCCGUUUUUUGUCCAAAGCGGAGGCAA 60

QY 239 CCCCGCGCTTGGATTAATAACT 260
DB 61 CCCCGCGUUGGAUUAUAAACU 82

RESULT 14

ADJ56742/c
ID ADJ56742 standard; DNA; 108 BP.

XX AC

XX ADJ56742;

XX 06-MAY-2004 (first entry)

XX Mutagenic PCR primer used to amplify hepatitis GB virus B cDNA SeqID 12.

XX ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
KW vaccine; virucidal; antiinflammatory; PCR; primer; mutagenic.

XX Hepatitis GB virus B.
OS Synthetic.

XX WO2004005498-A1.

XX 15-JAN-2004.

XX 02-JUL-2003; 2003WO-US021002.

XX 03-JUL-2002; 2002US-00189359.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (INSP) INST PASTEUR.

XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;

XX WPI; 2004-091362/09.

XX New chimeric GBV-B polynucleotide, useful as a model for hepatitis C PT virus, for identifying compounds active against a viral infection, or for PT developing hepatitis C virus preventive and therapeutic treatments.

XX Example 22; SEQ ID NO 12; 108pp; English.

XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV CC polynucleotides. Specifically, it refers to using the hepatotropic CC flavivirus GBV-B that has a unique phylogenetic relationship to the human CC hepatitis C virus (HCV) and can serve as a surrogate virus in drug CC discovery efforts related to antiviral drug development. The present CC invention describes the construction of an infectious molecular clone CC using the newly determined 3' terminal sequence of GBV-B. Furthermore, CC the GBV-B/HCV chimaeras exhibit liver-specific expression and express HCV CC envelope proteins such that they can have utility as a vaccine immunogen CC for hepatitis C. In addition, they can be used for screening compounds CC preventative and therapeutic treatments. Accordingly, these compositions CC exhibit virucidal, antiinflammatory and hepatotropic activities. This CC oligonucleotide sequence is a mutagenic PCR primer used to amplify the CC GBV-B cDNA sequence of the invention.

XX Sequence 108 BP; 23 A; 33 C; 30 G; 22 T; 0 U; 0 Other;

Query Match 25.4%; Score 66; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.7e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 254
DB 108 GTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 49

QY 255 AAAACT 260
|||||
Db 48 AAAACT 43

RESULT 15
ADJ64254/c
ID ADJ64254 standard; DNA; 108 BP.
XX
AC ADJ64254;
XX
XX 20-MAY-2004 (first entry)
XX
DE Hepatitis GB virus B infectious clone related primer seqid 12.
XX
XX antinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
KW hepatitis C virus; HCV; 3' terminal; primer; ss; PCR; infectious clone.
XX
XX
OS Hepatitis GB virus B.
XX
XX US2004039187-A1.
PN
XX 26-FEB-2004.
PD
XX
XX 03-JUL-2002; 2002US-00189359.
PF
XX
XX 04-JUN-1999; 99US-0137665P.
PR
XX 05-JUN-2000; 2000US-00587653.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
PA (INSP) INST PASTEUR.
PA
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
PI
XX WPI; 2004-203294/19.
DR
XX
XX New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
PT diagnosing and in treating HCV and in investigating the mechanisms for
PT the different biological properties of the viruses.
XX
XX Example 22; SEQ ID NO 12; 58pp; English.
PS
XX The invention describes a new isolated polynucleotide (I) encoding a 3'
CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
CC GBV-B genome, where at least part, but not all of a 5' nontranslated
CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
CC or chimaeras are useful diagnosing or treating hepatitis C virus (HCV)
CC and in investigating the mechanisms for the different biological
CC properties of the viruses. This sequence represents a primer used in the
CC construction of a Hepatitis GB virus B (GBV-B) infectious clone.
XX
XX Sequence 108 BP; 23 A; 33 C; 30 G; 22 T; 0 U; 0 Other;
SQ

Query Match 25.4%; Score 66; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.7e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GTACCAAGGGCTGCACCCCGGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 254
|||||
Db 108 GTACCAAGGGCTGCACCCCGGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 49

QY 255 AAAACT 260
|||||
Db 48 AAAACT 43

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 13:50:44 ; Search time 127 Seconds
(without alignments)
3639.103 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gaggttggcgaccatggtgg.....cccgcttggaattaaaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	258.4	99.4	309	US-09-579-302-1	Sequence 1, Appli
2	257.4	99.0	259	US-09-579-302-2	Sequence 2, Appli
C 3	257.4	99.0	259	US-09-579-302-3	Sequence 3, Appli
4	257.4	99.0	259	US-09-579-302-4	Sequence 4, Appli
C 5	257.4	99.0	259	US-09-579-302-5	Sequence 5, Appli
6	82	31.5	82	US-09-579-302-14	Sequence 14, Appl
7	30.8	11.8	133358	US-09-949-016-16364	Sequence 16964, A
8	30.8	11.8	133360	US-09-949-016-12651	Sequence 12651, A
9	30.4	11.7	20444	US-09-949-016-15750	Sequence 15750, A
C 10	29.8	11.5	601	US-09-949-016-182173	Sequence 182173,
C 11	29.8	11.5	24183	US-08-943-731-3	Sequence 3, Appli
C 12	29.4	11.3	1002	US-09-468-2538-53	Sequence 53, Appl
C 13	29.4	11.3	1005	US-09-468-2538-51	Sequence 51, Appl
C 14	29.4	11.3	1047	US-09-468-2538-61	Sequence 61, Appl
C 15	29.4	11.3	1047	US-09-468-2538-62	Sequence 62, Appl
C 16	29.4	11.3	2092	US-09-468-2538-13	Sequence 13, Appl
C 17	29.4	11.3	2473	US-09-468-2538-15	Sequence 15, Appl
C 18	29.2	11.2	1368	US-07-677-413-1	Sequence 1, Appli
19	29.2	11.2	1571	US-07-677-413-2	Sequence 2, Appli
20	29.2	11.2	1830121	US-09-557-884-1	Sequence 1, Appli
21	29.2	11.2	1830121	US-09-643-990A-1	Sequence 1, Appli
22	29.2	11.2	1830121	US-10-158-865-1	Sequence 1, Appli
23	29	11.2	1630	US-09-949-016-3210	Sequence 3210, Ap
24	29	11.2	1723	US-07-841-646-28	Sequence 28, Appl

25	29	11.2	1723	2	US-07-901-703-10	Sequence 10, Appl
26	29	11.2	1723	2	US-08-147-023-28	Sequence 28, Appl
27	29	11.2	1723	2	US-08-206-864-3	Sequence 3, Appli
28	29	11.2	1723	2	US-08-278-729A-20	Sequence 20, Appl
29	29	11.2	1723	2	US-08-480-528A-7	Sequence 7, Appli
30	29	11.2	1723	2	US-08-479-666-7	Sequence 7, Appli
31	29	11.2	1723	2	US-08-155-343A-20	Sequence 20, Appl
32	29	11.2	1723	2	US-08-408-672-20	Sequence 20, Appl
33	29	11.2	1723	2	US-08-643-563A-20	Sequence 20, Appl
34	29	11.2	1723	2	US-08-447-570-28	Sequence 20, Appl
35	29	11.2	1723	2	US-08-643-763A-20	Sequence 20, Appl
36	29	11.2	1723	2	US-08-462-623-20	Sequence 20, Appl
37	29	11.2	1723	2	US-08-451-953A-20	Sequence 20, Appl
38	29	11.2	1723	2	US-08-459-346-5	Sequence 5, Appli
39	29	11.2	1723	2	US-08-445-468A-20	Sequence 20, Appl
40	29	11.2	1723	2	US-08-901-200A-7	Sequence 7, Appli
41	29	11.2	1723	2	US-08-449-700-28	Sequence 28, Appl
42	29	11.2	1723	2	US-08-449-699A-28	Sequence 28, Appl
43	29	11.2	1723	2	US-08-461-397A-20	Sequence 20, Appl
44	29	11.2	1723	2	US-08-912-088-20	Sequence 20, Appl
45	29	11.2	1723	3	US-08-278-730A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-579-302-1
; Sequence 1, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 309
; TYPE: DNA
; ORGANISM: GBV-B-like virus
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1)..(309)
; OTHER INFORMATION: cdna complementary to the last 49 published 3'UTR
; OTHER INFORMATION: nucleotides of GBV-B plus a novel nucleotide (C)
; OTHER INFORMATION: in position 44 and 259 novel nucleotide sequence
; OTHER INFORMATION: at the 3' end.
US-09-579-302-1

Query Match 99.4%; Score 258.4; DB 3; Length 309;
Best Local Similarity 99.6%; Pred. No. 8.2e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAGTTTGGCGACCATGGTGCATCAGAACCGTTTCGGGTGAAGCCATGCTGAAGGGGAT 60
Db 50 GAGTTTGGCGACCATGGTGCATCAGAACCGTTTCGGGTGAAGCCATGCTGAAGGGGAT 109
Qy 61 GAGTCCTCTTGGCTCATCCACAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTG 120
Db 110 GAGTCCTCTTGGCTCATCCACAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTG 169
Qy 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGCACGCTTACGACGATTTGTC 180
Db 170 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGCACGCTTACGACGATTTGTC 229
Qy 181 CTGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTTCAGAGGAGGCAACC 240
Db 230 CTGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTTCAGAGGAGGCAACC 289

```
QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 290 CCCGCTTGGAAATTAATAAACT 309

RESULT 2
US-09-579-302-2
; Sequence 2, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 259
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-2

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 76.1%; Pred. No. 1.7e-80;
Matches 197; Conservative 61; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
Db 1 AGUUGGCGACCAUGUGUAUCAGAACCGUUUUGGGUGAAGCAUGGUCUGAAGGGGAUG 60

QY 62 ACGTCCCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTGT 121
Db 61 ACGUCCCUUGGUCUACCAACAAAAACCGUCUCGGUGGUGAGGAGUCCUGGUGUGU 120

QY 122 GGGAGCAGTCAGTATAATTCCTCGTGTGTGTGAGCGCTCAGACGTATTGTTCGCG 181
Db 121 GGGAGCAGUCAGUAUAUUCGUGUGUGUGAGCGCCUCACGACGUAUCUUGUCCGC 180

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTGTTCGAAGCGGAGGCAACCC 241
Db 181 UUGUCAGAGCGUAGUACCAAGGGUCACCCCGUUUUUUUGUCCAAAGCGGAGGCAACCC 240

QY 242 CCGCTTGGAAATTAATAAACT 260
Db 241 CCGCUUGGAUUUAAAAACU 259

RESULT 3
US-09-579-302-3/c
; Sequence 3, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 259
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-3

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 290 CCCGCTTGGAAATTAATAAACT 309

RESULT 4
US-09-579-302-4
; Sequence 4, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: DNA
; ORGANISM: GBV-B-like virus
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (1)..(259)
; OTHER INFORMATION: cDNA of positive strand 3'UTR of GBV-B
US-09-579-302-4

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
Db 1 AGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60

QY 62 ACGTCCCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTGT 121
Db 61 ACGTCCCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTGT 120

QY 122 GGGAGCAGTCAGTATAATTCCTCGTGTGTGTGAGCGCTCAGACGTATTGTTCGCG 181
Db 121 GGGAGCAGTCAGTATAATTCCTCGTGTGTGTGAGCGCTCAGACGTATTGTTCGCG 180

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTGTTCGAAGCGGAGGCAACCC 241
Db 181 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTGTTCGAAGCGGAGGCAACCC 240

QY 242 CCGCTTGGAAATTAATAAACT 260
Db 241 CCGCTTGGAAATTAATAAACT 259

RESULT 5
US-09-579-302-5/c
```

Sequence 5, Application US/09579302
Patent No. 6627437
GENERAL INFORMATION:
APPLICANT: Traboni, Cinzia
TITLE OF INVENTION: No. 6627437el GBV sequence
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/579,302
CURRENT FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: GB 9912432.3
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 259
TYPE: DNA
ORGANISM: GBV-B-like virus
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (1)..(259)
OTHER INFORMATION: cdna of negative strand 3'UTR of GBV-B
US-09-579-302-5

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTTTCGGCAGCATGGTGGATCAGAACCGCTTTTCGGGTGAAGCCATGCTCAAGGGGATG 61
Db 259 AGTTTCGGCAGCATGGTGGATCAGAACCGCTTTTCGGGTGAAGCCATGCTCAAGGGGATG 200

Qy 62 ACCTCCCTCTGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 121
Db 199 ACCTCCCTCTGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 140

Qy 122 GCGAAGCAGTCAGTAAATTCCTCGTGTGTGTGAGCCCTCAGCAGTATTGTCTCCG 181
Db 139 GCGAAGCAGTCAGTAAATTCCTCGTGTGTGTGAGCCCTCAGCAGTACTTGTCTCCG 80

Qy 182 TGTGCGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
Db 79 TGTGCGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTGTTCGAAGCGGAGGCAACCC 20

Qy 242 CCGCTTGGAAATTAATAACT 260
Db 19 CCGCTTGGAAATTAATAACT 1

RESULT 6
US-09-579-302-14
Sequence 14, Application US/09579302
Patent No. 6627437
GENERAL INFORMATION:
APPLICANT: Traboni, Cinzia
TITLE OF INVENTION: No. 6627437el GBV sequence
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/579,302
CURRENT FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: GB 9912432.3
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 82
TYPE: RNA
ORGANISM: GBV-B-like virus
US-09-579-302-14

Query Match 31.5%; Score 82; DB 3; Length 82;
Best Local Similarity 79.3%; Pred. No. 5.2e-19;
Matches 65; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 179 CGCTGTGCAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCA 238
Db 179 CGCTGTGCAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCA 238

Sequence 5, Application US/09579302
Patent No. 6627437
GENERAL INFORMATION:
APPLICANT: Traboni, Cinzia
TITLE OF INVENTION: No. 6627437el GBV sequence
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/579,302
CURRENT FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: GB 9912432.3
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 259
TYPE: DNA
ORGANISM: GBV-B-like virus
FEATURE:
NAME/KEY: 3'UTR
LOCATION: (1)..(259)
OTHER INFORMATION: cdna of negative strand 3'UTR of GBV-B
US-09-579-302-5

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGTTTCGGCAGCATGGTGGATCAGAACCGCTTTTCGGGTGAAGCCATGCTCAAGGGGATG 61
Db 259 AGTTTCGGCAGCATGGTGGATCAGAACCGCTTTTCGGGTGAAGCCATGCTCAAGGGGATG 200

Qy 62 ACCTCCCTCTGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 121
Db 199 ACCTCCCTCTGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTTGGCTGTGT 140

Qy 122 GCGAAGCAGTCAGTAAATTCCTCGTGTGTGTGAGCCCTCAGCAGTATTGTCTCCG 181
Db 139 GCGAAGCAGTCAGTAAATTCCTCGTGTGTGTGAGCCCTCAGCAGTACTTGTCTCCG 80

Qy 182 TGTGCGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
Db 79 TGTGCGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTGTTCGAAGCGGAGGCAACCC 20

Qy 242 CCGCTTGGAAATTAATAACT 260
Db 19 CCGCTTGGAAATTAATAACT 1

RESULT 6
US-09-579-302-14
Sequence 14, Application US/09579302
Patent No. 6627437
GENERAL INFORMATION:
APPLICANT: Traboni, Cinzia
TITLE OF INVENTION: No. 6627437el GBV sequence
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/579,302
CURRENT FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: GB 9912432.3
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 82
TYPE: RNA
ORGANISM: GBV-B-like virus
US-09-579-302-14

Query Match 31.5%; Score 82; DB 3; Length 82;
Best Local Similarity 79.3%; Pred. No. 5.2e-19;
Matches 65; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

Qy 179 CGCTGTGCAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCA 238
Db 179 CGCTGTGCAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCA 238

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-3

Query Match 11.5%; Score 29.8; DB 3; Length 24183;
Best Local Similarity 54.0%; Pred. No. 12;
Matches 61; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 121 TGGGAGCAGTCAGTATATCCCGTCGTCGTGGTGACGCCCTCAGACGTATTTGTCGCG 180
Db |||||
4413 TTGAGGTAGTCAGGGGACTCACCGTTATCTGGCTGGCAGCTCATGGCAAGTTCTCTC 4354

QY 181 CTGTGCAGACGCTAGTACCAAGGCTGCACCCCGTTTTTGTTCCTCAAGCGGAG 233
Db |||||
4353 CTGGCCGCGAGGGGCTCACAATGGATCAGCATCCATTGAATCAACTGGAG 4301

RESULT 12
US-09-468-253B-53/c
; Sequence 53, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-468-253B-53

Query Match 11.3%; Score 29.4; DB 3; Length 1002;
Best Local Similarity 50.2%; Pred. No. 4.5;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;
QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGCTCCCTCTGCGCTCATCCACAAAAACCGTCTC 95
Db 900 GGGGTATTCATAGGGTGTAGATTAAACCAGCTTCAGGCCCTGGAGGACTATTGCAGC 841
QY 96 GGGTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 781
Db 840 AGTTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 781
QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 780 TATGGTAGGGCCACGCTGGCGTAGGTGTACGACAGCAGCTGGTGGGAGGACAGGCGCAGG 721

QY 211 CCGGGTTTTTGTTCCTCAAGCGGAG 233
Db 720 CCAGTGATGATCCTTGGAGCAG 698
Query Match 11.3%; Score 29.4; DB 3; Length 1047;
Best Local Similarity 50.2%; Pred. No. 4.6;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;
QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGCTCCCTCTGCGCTCATCCACAAAAACCGTCTC 95
Db 878 GGGGTACTCATAGGGTGTATAGATTAAACCAGCTTCGGGCCCTGGAGGAACTATTGCAGC 819
QY 96 GGGTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 759
Db 818 AGTTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 759
QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 758 TATGGTAGGGCCAGCTGGCGTAGGAGTACGCGAGGCGAGCTGGTGGGAGAACCCGCGCAGG 699

RESULT 13
US-09-468-253B-51/c
; Sequence 51, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127

;
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-468-253B-51

Query Match 11.3%; Score 29.4; DB 3; Length 1005;
Best Local Similarity 50.2%; Pred. No. 4.5;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGCTCCCTCTGCGCTCATCCACAAAAACCGTCTC 95
Db 900 GGGGTACTCATAGGGTGTATAGATTAAACCAGCTTCGGGCCCTGGAGGAACTATTGCAGC 841

QY 96 GGGTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 781
Db 840 AGTTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 781

QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 780 TATGGTAGGGCCAGCTGGCGTAGGAGTACGCGAGGCGAGCTGGTGGGAGAACCCGCGCAGG 721

QY 211 CCGGGTTTTTGTTCCTCAAGCGGAG 233
Db 720 CCAGTGATGATCCTTGGAGCAG 698

RESULT 14
US-09-468-253B-61/c
; Sequence 61, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-468-253B-61

Query Match 11.3%; Score 29.4; DB 3; Length 1047;
Best Local Similarity 50.2%; Pred. No. 4.6;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGCTCCCTCTGCGCTCATCCACAAAAACCGTCTC 95
Db 878 GGGGTACTCATAGGGTGTATAGATTAAACCAGCTTCGGGCCCTGGAGGAACTATTGCAGC 819

QY 96 GGGTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 759
Db 818 AGTTGGGTGAGGAGTCTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAGGCAT 759

QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 758 TATGGTAGGGCCAGCTGGCGTAGGAGTACGCGAGGCGAGCTGGTGGGAGAACCCGCGCAGG 699

QY 211 CCGGGTTTTTGTTCCTCAAGCGGAG 233
Db 698 CCAGTGATGATCCTTGGAGCAG 676

RESULT 15

```

US-09-468-253B-62/c
; Sequence 62, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; TITLE OF INVENTION: Differential Gene Expression
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-468-253B-62

Query Match      11.3%; Score 29.4; DB 3; Length 1047;
Best Local Similarity 50.2%; Pred. No. 4.6;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY      36  GGTGAAGCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCACAAAACCGTCTC 95
Db      878  GGGTATTTCATAGGTGTGTAGATTAAACCAGCTTCAGGCCCTGGAGGACTATTGCAGC 819

QY      96  GGGTCGGGTGAGGAGTCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTC-----GTG 150
Db      818  AGTTGGGTGAGGAGTTCGGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAAGGCAT 759

QY     151  TGTGTGAGCCCTCAGGAGTATTGTCCGCTGTGAGAGCGTAGTACCAAGGGCTGCAC 210
Db     758  TATGTTAGGGCCAGCTGGCGTAGGTGTACGAGAGCAGCTGGTGGGAGGACAGGCGCAGG 699

QY     211  CCCGGTTTTTTTCCAAAGCGGAG 233
Db     698  CCCAGTGATGATCCTTGGAGCAG 676

```

Search completed: December 7, 2005, 19:45:11
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 17:19:59 ; Search time 535 Seconds
(without alignments)
4018.763 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gaggttggcgaccatgtggg.....cccgcttggaattaaaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	260	100.0	9399	3 US-09-742-659-1 Sequence 1, Appli
2	260	100.0	9399	7 US-10-189-359-2 Sequence 2, Appli
3	258.4	99.4	8069	9 US-10-501-412-1 Sequence 1, Appli
4	258.4	99.4	9397	9 US-10-501-412-2 Sequence 2, Appli
5	254.4	97.8	260	7 US-10-189-359-1 Sequence 1, Appli
C 6	66	25.4	108	7 US-10-189-359-12 Sequence 12, Appli
C 7	32	12.3	3179	6 US-10-168-651-36 Sequence 36, Appli
8	31.8	12.2	645	5 US-10-027-632-132082 Sequence 132082,
9	31.8	12.2	645	6 US-10-027-632-132082 Sequence 132082,
10	31.4	12.1	645	4 US-09-925-065A-425966 Sequence 425966,
11	31	11.9	1095	6 US-10-000-897-36 Sequence 36, Appli
12	31	11.9	1095	8 US-10-818-168-36 Sequence 36, Appli
13	30.6	11.8	1150	4 US-09-925-065A-551953 Sequence 551953,
14	30.6	11.8	3772	10 US-11-097-143-31040 Sequence 31040, A
15	30.2	11.6	645	4 US-09-925-065A-425967 Sequence 425967,
16	30	11.5	381	8 US-10-425-115-30124 Sequence 30124, A
17	30	11.5	626	4 US-09-925-065A-487889 Sequence 487889,
18	30	11.5	2499	7 US-10-267-502-86 Sequence 86, Appli
19	30	11.5	3855	10 US-11-097-143-42398 Sequence 42398, A
20	30	11.5	4497	10 US-11-097-143-54233 Sequence 54233, Ap
C 21	30	11.5	10907	10 US-11-097-143-42397 Sequence 42397, A
C 22	30	11.5	13815	10 US-11-097-143-54222 Sequence 54222, Ap
C 23	29.8	11.5	600	9 US-10-972-079-75557 Sequence 75557, A

ALIGNMENTS

RESULT 1

US-09-742-659-1
; Sequence 1, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9399
; TYPE: DNA
; ORGANISM: GB virus-B
US-09-742-659-1

Query Match 100.0%; Score 260; DB 3; Length 9399;
Best Local Similarity 100.0%; Pred. No. 6.2e-82;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9140 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAAGCCATGGTCTGAAGGGGAT 9199
QY 61 GAGCTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
DB 9200 GAGCTCCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 9259
QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCAGACGATATTTGTCCG 180
DB 9260 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGTGAGCGCTCAGACGATATTTGTCCG 9319
QY 181 CTGTGCAGACGCTAGTACCAAGGGCTGCCCGGTTTTTTTGTTCGAAGGGGGGCAACC 240
DB 9320 CTGTGCAGACGCTAGTACCAAGGGCTGCCCGGTTTTTTTGTTCGAAGGGGGGCAACC 9379

```
QY 241 CCCGCTTGGAAATTAATAACT 260
Db 9380 CCCGCTTGGAAATTAATAACT 9399

RESULT 2
US-10-189-359-2
; Sequence 2, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Annette
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: .10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9399
; TYPE: DNA
; ORGANISM: GBV-A-like virus
US-10-189-359-2

Query Match 100.0%; Score 260; DB 7; Length 9399;
Best Local Similarity 100.0%; Pred. No. 6.2e-82;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 9140 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 9199

QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
Db 9200 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 9259

QY 121 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTGTGAGCGCTCAGCGTATTTGTCCG 180
Db 9260 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTGTGAGCGCTCAGCGTATTTGTCCG 9319

QY 181 CTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACC 240
Db 9320 CTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACC 9379

QY 241 CCCGCTTGGAAATTAATAACT 260
Db 9380 CCCGCTTGGAAATTAATAACT 9399

RESULT 3
US-10-501-412-1
; Sequence 1, Application US/10501412
; Publication No. US20050239205A1
; GENERAL INFORMATION:
; APPLICANT: De Tomasi, Amedeo
; APPLICANT: Graziani, Rita
; APPLICANT: Paonessa, Giacomo
; APPLICANT: Tranboni, Cinzia
; TITLE OF INVENTION: GB VIRUS B BASED REPLICONS AND REPLICON
; FILE REFERENCE: ITR0037YP
; CURRENT APPLICATION NUMBER: US/10/501,412
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: PCT/EP03/00281
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,655
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/348,573
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 9397
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GBV-B Replicon
US-10-501-412-2

Query Match 99.4%; Score 258.4; DB 9; Length 9397;
Best Local Similarity 99.6%; Pred. No. 2.3e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 9138 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 9197

QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 8069
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GBV-B Replicon
US-10-501-412-1

Query Match 99.4%; Score 258.4; DB 9; Length 8069;
Best Local Similarity 76.2%; Pred. No. 2.2e-81;
Matches 198; Conservative 61; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 7810 GAGUUGGCGACCAUGGUGAUCAGCCGUUUUGGGUGAGCCGAUGGUCUGAAGGGGAU 7869

QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
Db 7870 GAGUUGGCGACCAUGGUGAUCAGCCGUUUUGGGUGAGCCGAUGGUCUGAAGGGGAU 7929

QY 121 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTGTGAGCGCTCAGCGTATTTGTCCG 180
Db 7930 UGGGAAGCAGUCAGUAUAUUCGCGUCUGUGUGAGCGCCUCACGACGACUUGUCG 7989

QY 181 CTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACC 240
Db 7990 CUGUGCAGAGCGUAGUACCAAGGCGUGACCCCGGUUUUUUGUCCAAGCGGAGGCAACC 8049

QY 241 CCCGCTTGGAAATTAATAACT 260
Db 8050 CCCGCUUGGAUUUAAAACU 8069

RESULT 4
US-10-501-412-2
; Sequence 2, Application US/10501412
; Publication No. US20050239205A1
; GENERAL INFORMATION:
; APPLICANT: De Tomasi, Amedeo
; APPLICANT: Graziani, Rita
; APPLICANT: Paonessa, Giacomo
; APPLICANT: Tranboni, Cinzia
; TITLE OF INVENTION: GB VIRUS B BASED REPLICONS AND REPLICON
; FILE REFERENCE: ITR0037YP
; CURRENT APPLICATION NUMBER: US/10/501,412
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: PCT/EP03/00281
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,655
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/348,573
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 9397
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GBV-B Replicon
US-10-501-412-2

Query Match 99.4%; Score 258.4; DB 9; Length 9397;
Best Local Similarity 99.6%; Pred. No. 2.3e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 9138 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 9197

QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
```

Db 9198 GACGTCCTTCTGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 9257
Qy 121 TGGGAAGCAGTCAGTATATTCCTCGTCTGTGTGTGTGACCGCTCACGACGATTTGTGTCG 180
Db 9258 TGGGAAGCAGTCAGTATATTCCTCGTCTGTGTGTGTGACCGCTCACGACGATTTGTGTCG 9317
Qy 181 CTGTGCAGCGGTAGTATCAAGGGCTGCACCCCGGTTTTTTGTTTCAAGCGGAGGCAACC 240
Db 9318 CTGTGCAGCGGTAGTATCAAGGGCTGCACCCCGGTTTTTTGTTTCAAGCGGAGGCAACC 9377
Qy 241 CCCGCTTGAATTAATAACT 260
Db 9378 CCCGCTTGAATTAATAACT 9397

RESULT 5
US-10-189-359-1
; Sequence 1, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: SANGAR, Annette
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 260
; TYPE: RNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: GB VIRUS B
US-10-189-359-1

Query Match 97.8%; Score 254.4; DB 7; Length 260;
Best Local Similarity 75.8%; Pred. No. 2.6e-80;
Matches 194; Conservative 61; Mismatches 1; Indels 0; Gaps 0;
Qy 5 TTGGCGACCATGTCATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATGACG 64
Db 4 UUGCGACCAUGGUGUACAGAACCGUUUCGGGGAAGCCAUUGGUGUAGGAGGUGAGCG 63
Qy 65 TCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTGTGGG 124
Db 64 UCCUUCUGGUCUACCAAAAAACGUCUCGGGUGGUGAGGAGUCCUGGUGUGGG 123
Qy 125 AAGCAGTCAGTATATTCCTCGTCTGTGTGTGTGACCGCTCACGACGATTTGTGTCGCTGT 184
Db 124 AAGCAGUCAGUAUAUUCGUGUGUGUGAGCCUCACGACGUAUUGUCGCGUGU 183
Qy 185 GCAGCGGTAGTATCAAGGGCTGCACCCCGGTTTTTTGTTTCAAGCGGAGGCAACCCCG 244
Db 184 GCAGCGGAGUACCAAGGGGUGCACC CGGUGUUUUUUUUUCCAAAGCGGAGGCAACCCCG 243
Qy 245 CTGGAATTAATAACT 260
Db 244 CUUGGAUUUAAAAACU 259

RESULT 6
US-10-189-359-12/c
; Sequence 12, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Annette
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.

; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-189-359-12

Query Match 25.4%; Score 66; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 195 GTACCAAGGGCTGCACCCCGGTTTTTTGTTTCAAGCGGAGGCAACCCCGCTTGGAAATTA 254
Db 108 GTACCAAGGGCTGCACCCCGGTTTTTTGTTTCAAGCGGAGGCAACCCCGCTTGGAAATTA 49
Qy 255 AAAACT 260
Db 48 AAAACT 43

RESULT 7
US-10-168-651-36/c
; Sequence 36, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameen R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
; 60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 3179
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 3232992CBI
US-10-168-651-36

Query Match 12.3%; Score 32; DB 6; Length 3179;
Best Local Similarity 50.7%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

```
QY 10 GACCATGCTGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATGACGTCCCT 69
Db 2836 GAGCATTTGCCAAGCTGCTCTTGGGGGTCCCCCATTTGGGCACAAAGCCCTCACCCCCCA 2777
QY 70 TCTGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGATCCTTGGCTGTGTGGGAAGCA 129
Db 2776 CCCCACCCCGCACCAAGAACCCCTTGTATTCAGCAGAAAGTCAGGCAGTGGGTGACGCA 2717
QY 130 GTCAGTATAATTCCTCGTCTGTGTGTGACGC 161
Db 2716 GGGGGCTAAGCTCCCCCAAGCGGTTGGGC 2685

RESULT 8
US-10-027-632-132082
; Sequence 132082, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132082
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-132082

Query Match 12.2%; Score 31.8; DB 5; Length 645;
Best Local Similarity 54.8%; Pred. No. 1.2;
Matches 63; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 41 AGCCATGCTTGAAGGGATGACGTCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTG 100
Db 417 ACCAAGGAACCTTAGGAGGTGAAGTCATTTCTGGGTGGAGATCATCCCTGTCAAGGGTG 476
QY 101 GGTGAGGAGTCTGCTGTGTGGGAAGCAGTCAGTATAATTCCTCGTGTGTGG 155
Db 477 GGTGGGAGATCATCCCGTCAAGGGTGGTGGAGATCATCCCGTCAAGGGTGG 531

RESULT 9
US-10-027-632-132082
; Sequence 132082, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132082
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-132082

Query Match 12.2%; Score 31.8; DB 6; Length 645;
Best Local Similarity 54.8%; Pred. No. 1.2;
Matches 63; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 41 AGCCATGCTTGAAGGGATGACGTCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTG 100
Db 417 ACCAAGGAACCTTAGGAGGTGAAGTCATTTCTGGGTGGAGATCATCCCTGTCAAGGGTG 476
QY 101 GGTGAGGAGTCTGCTGTGTGGGAAGCAGTCAGTATAATTCCTCGTGTGTGG 155
Db 477 GGTGGGAGATCATCCCGTCAAGGGTGGTGGAGATCATCCCGTCAAGGGTGG 531

RESULT 10
US-09-925-065A-425966
; Sequence 425966, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425966
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-425966

Query Match 12.1%; Score 31.4; DB 4; Length 645;
Best Local Similarity 52.8%; Pred. No. 1.6;
Matches 65; Conservative 1; Mismatches 57; Indels 0; Gaps 0;

QY 105 AGGATCTCTGCTGTGTGGGAAGCAGTCAGTATAATTCCTCGTGTGTGGTGAAGCCTC 164
Db 2 AGGCACCTTGAACCTGAGCCACCGCTCCAGTGGAACTGCAAGTGGTTATGCTC 61
QY 165 ACAGCGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGTGCACCCCGGTTTTTGTTC 224
Db 62 GCCAAGTGTGTGGGATCTGTGTGGGGCTTACCATCACCTGTACTTCCCTCTCCCTGTGC 121
```

```
QY 225 CAA 227
Db 122 TAA 124

RESULT 11
US-10-000-897-36
; Sequence 36, Application US/10000897
; Publication No. US20030165852A1
; GENERAL INFORMATION:
; APPLICANT: Schueler, P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE
; FILE OF INVENTION: MATERNAL CIRCULATION
; FILE REFERENCE: 11012-004-999
; CURRENT APPLICATION NUMBER: US/10/000,897
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/248,882
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-000-897-36

Query Match 11.9%; Score 31; DB 6; Length 1095;
Best Local Similarity 47.8%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;
Matches 91; Conservative 0;

QY 4 TTTCGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGAC 63
Db 620 TTTCGACAAATGCTCATTAATATATTTTCTATTCTAGACCATATCTGAAGAAATAAC 679

QY 64 GTCCCTTCGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTCGGCTGTGTGG 123
Db 680 ATTTTAAAGGCTCTACCATAGACATATCTAGTAATGTGTGTGTGTGTGTGTGTGT 739

QY 124 GAAGCATGATGATTAATCCCGTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
Db 740 GTGTATGTATGTATAGTTCGGGAGAGGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 799

QY 184 TGCAGAGCGTA 194
Db 800 GGAGGACTGGA 810

RESULT 12
US-10-818-168-36
; Sequence 36, Application US/10818168
; Publication No. US20040185495A1
; GENERAL INFORMATION:
; APPLICANT: Schueler, P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE MAT
; FILE OF INVENTION: CIRCULATION
; FILE REFERENCE: 11012-004-999
; CURRENT APPLICATION NUMBER: US/10/818,168
; CURRENT FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: 60/248,882
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-818-168-36

Query Match 11.9%; Score 31; DB 8; Length 1095;
Best Local Similarity 47.6%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;
Matches 91; Conservative 0;

QY 4 TTTCGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGAC 63
```

```
Db 620 TTTCGACAAATGCTCATTAATATATTTTCTATTCTAGACCATATCTGAAGAAATAAC 679
QY 64 GTCCCTTCGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTCGGCTGTGTGG 123
Db 680 ATTTTAAAGGCTCTACCATAGACATATCTAGTAATGTGTGTGTGTGTGTGTGTGT 739

QY 124 GAAGCATGATGATTAATCCCGTCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
Db 740 GTGTATGTATGTATAGTTCGGGAGAGGATAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 799

QY 184 TGCAGAGCGTA 194
Db 800 GGAGGACTGGA 810

RESULT 13
US-09-925-065A-551953
; Sequence 551953, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 551953
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-551953

Query Match 11.8%; Score 30.6; DB 4; Length 1150;
Best Local Similarity 58.1%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 54; Conservative 0;

QY 8 GCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACGTCC 67
Db 998 GTGCACATGGAGCCTTAGCTCTGATTCAGTTCCAGTTCAGCAAAATTTATGAAGAAGATGATTCCC 1057

QY 68 CTTCTGGCTCATCCACAAAACCGTCTCGGGTG 100
Db 1058 CATATGGATGACAGATAAACACCCCTGTGGCGAG 1090

RESULT 14
US-11-097-143-31040
; Sequence 31040, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; FILE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
```

; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31040
; LENGTH: 3772
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-31040

Query Match 11.8%; Score 30.6; DB 10; Length 3772;
Best Local Similarity 58.1%; Pred. No. 4.9;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 54 AGGGATGACGTCCTCTGGCTCATCCAAAAACCCGTCCTCGGGTGGGTGAGGAGTCCT 113
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 73 AGCGCTGAAGTCCTGGCTGACTGAACGAACGCCGTTCCGGCACGTTGTGAGTTGG 132
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 114 GGCTGTGGGAAGCAGTCAGTATATATCCCGT 146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 133 CTCGTAGAGTATGGCAGTCCTGGAGTTCCCT 165
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-925-065A-425967
; Sequence 425967, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425967
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-425967

Query Match 11.6%; Score 30.2; DB 4; Length 645;
Best Local Similarity 52.8%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 105 AGGAGTCCTGGGTGTGGGAAGCAGTCAGTATATATCCGTCGTGTGGTGAGCCCTC 164
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 AGGCCACCTGGAACCTGACCCACCGCTGCCAGTGGAAACACTGCACGAGTGGTTATGCCTC. 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 165 ACGACGTATTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTC 224
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 62 GCCAAGTGTGTGGGATCTGTGTGGGGCTTACCATCACCTGCTACTCCCTCTCCCTGTGC 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 225 CAA 227
|||
DB 122 TAA 124
|||

Search completed: December 7, 2005, 20:59:42
Job time : 539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 19:43:05 ; Search time 146 Seconds
(without alignments)
665.769 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gsgtttggcgaccatgtgg.....ccccgttggaataaaact 260

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues
Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New.*
1: /cgn2_6/prodata/2/pubnpa/US09_NEW_PUB.seq:
2: /cgn2_6/prodata/2/pubnpa/US06_NEW_PUB.seq:
3: /cgn2_6/prodata/2/pubnpa/US07_NEW_PUB.seq:
4: /cgn2_6/prodata/2/pubnpa/US08_NEW_PUB.seq:
5: /cgn2_6/prodata/2/pubnpa/PCT_NEW_PUB.seq:
6: /cgn2_6/prodata/2/pubnpa/US10_NEW_PUB.seq:
7: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq:
8: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq2:
9: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq3:
10: /cgn2_6/prodata/2/pubnpa/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	11.2	1723	7	US-11-051-568-28
2	28.6	11.0	1660	6	US-10-750-185-53296
3	28.6	11.0	319608	7	US-11-145-703-1
4	27.4	10.5	38703	7	US-11-052-544-28
5	27.2	10.5	1926	7	US-11-051-568-26
6	27.2	10.5	162537	7	US-11-121-086-59
7	26.6	10.2	169725	7	US-11-121-086-63
8	26.6	10.2	200628	7	US-11-121-086-62
9	26.4	10.2	1683	6	US-10-750-185-58438
10	26.4	10.2	2078	6	US-10-750-185-32595
11	26.4	10.2	2339	7	US-11-102-240-109
12	26.2	10.1	1267	6	US-10-750-185-39628
13	26.2	10.1	1431	6	US-10-750-185-39518
14	26	10.0	600	6	US-10-750-185-2422
15	26	10.0	1032	6	US-10-750-185-24904
16	26	10.0	3223	6	US-10-750-185-46677
17	26	10.0	167891	7	US-11-121-086-14
18	25.8	9.9	339	7	US-11-060-659-5
19	25.8	9.9	340	7	US-11-060-659-6
20	25.8	9.9	1413	6	US-10-750-185-37592
21	25.8	9.9	4539	6	US-10-750-185-56550
22	25.6	9.8	942	6	US-10-750-185-36167
23	25.6	9.8	1699	6	US-10-750-185-56336

C 24	25.6	9.8	2876	6	US-10-750-185-59433
C 25	25.6	9.8	128552	7	US-11-121-086-1
C 26	25.6	9.8	166639	7	US-11-121-086-52
C 27	25.4	9.8	797	6	US-10-750-185-36568
C 28	25.4	9.8	976	6	US-10-750-185-29663
C 29	25.4	9.8	1064	6	US-10-750-185-45839
C 30	25.4	9.8	1362	6	US-10-750-185-51469
C 31	25.4	9.8	1459	6	US-10-750-185-47911
C 32	25.4	9.8	1576	6	US-10-750-185-55314
C 33	25.4	9.8	1679	6	US-10-750-185-47978
C 34	25.4	9.8	13672	7	US-11-055-035-2
C 35	25.4	9.8	154548	7	US-11-121-086-33
C 36	25.2	9.7	540	6	US-10-750-185-25574
C 37	25.2	9.7	1479	6	US-10-467-9628-50
C 38	25.2	9.7	2253	6	US-10-750-185-49286
C 39	25.2	9.7	7008	6	US-10-821-234-758
C 40	25.2	9.7	169725	7	US-11-121-086-63
C 41	25.2	9.7	200628	7	US-11-121-086-62
C 42	25	9.6	690	6	US-10-750-185-29148
C 43	25	9.6	1516	6	US-10-750-185-53256
C 44	25	9.6	1673	6	US-10-750-185-32424
C 45	25	9.6	2307	6	US-10-750-185-29133

ALIGNMENTS

RESULT 1
US-11-051-568-28
; Sequence 28, Application US/11051568
; Publication No. US20050255141A1
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; OZKAYNAK, ENGIN
; KUBERASAMPATH, THANGAVEL
; RUEGER, DAVID C.
; PANG, ROY H.L.
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/11/051,568
; FILING DATE: 04-Feb-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 10/321,799
; FILING DATE: 17-DEC-2002
; APPLICATION NUMBER: US 09/148,925
; FILING DATE: 8-SEP-1998
; APPLICATION NUMBER: US 08/449,699
; FILING DATE: 24-MAY-1995
; APPLICATION NUMBER: US 08/147,023
; FILING DATE: 1-NOV-1993
; APPLICATION NUMBER: US 07/841,646
; FILING DATE: 21-FEB-1992
; APPLICATION NUMBER: US 07/827,052
; FILING DATE: 28-JAN-1992
; APPLICATION NUMBER: US 07/579,865
; FILING DATE: 7-SEP-1990
; APPLICATION NUMBER: US 07/621,849
; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/621,988

```
; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/810,560
; FILING DATE: 20-DEC-1991
; APPLICATION NUMBER: US 07/569,920
; FILING DATE: 20-AUG-1990
; APPLICATION NUMBER: US 07/600,024
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/599,543
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/616,374
; FILING DATE: 21-NOV-1990
; APPLICATION NUMBER: US 07/483,913
; FILING DATE: 22-FEB-1990
; APPLICATION NUMBER: US 07/179,406
; FILING DATE: 18-APR-1988
; APPLICATION NUMBER: US 07/232,630
; FILING DATE: 15-AUG-1988
; APPLICATION NUMBER: US 07/315,342
; FILING DATE: 23-FEB-1989
; APPLICATION NUMBER: US 07/660,162
; FILING DATE: 22-FEB-1991
; APPLICATION NUMBER: US 07/422,699
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,613
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,623
; FILING DATE: 17-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DIANA M. STEEL
; REGISTRATION NUMBER: 43,153
; REFERENCE/DOCKET NUMBER: STK-001CP6C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: HIPPOCAMPUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 490..1696
; OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
; /product= "hOP2-PP"
; /note= "hOP2 (cDNA)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-11-051-568-28

Query Match 11.2%; Score 29; DB 7; Length 1723;
Best Local Similarity 55.4%; Pred. No. 1.4; Mismatches 45; Indels 0; Gaps 0;
Matches 56; Conservative 0;

QY 1 GAGTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 807 GAGCTTCGTTACATGGTGGAGGAGACCGTGCCCTGGGCCACAGGAGCCCATTTGGAA 866

QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGG 101
DB 867 GGAGTTCCGCTTTGACCTGACCCAGATCCCGGCTGGGAGG 907

RESULT 2
US-10-185-53296/c
; Sequence 53296, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
```

```
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53296
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Bovine 19866880574203
US-10-750-185-53296

Query Match 11.0%; Score 28.6; DB 6; Length 1660;
Best Local Similarity 64.2%; Pred. No. 1.9; Mismatches 24; Indels 0; Gaps 0;
Matches 43; Conservative 0;

QY 184 TGCAGAGCGTAGTACCAAGGGTGCACCCCGGTTTTTGTTCACAGCGAGGCGCAACCCCC 243
DB 84 TGCAGAGCGGAGCAGCTCATGCTACAGCCAGATCTAGTTCAACCCGGAATCCACCACC 25

QY 244 GCTTGGGA 250
DB 24 TGCTGGA 18

RESULT 3
US-11-145-703-1
; Sequence 1, Application US/11145703
; Publication No. US20050260667A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihaïn, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: 53-US16.DIV
; CURRENT APPLICATION NUMBER: US/11/145,703
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US/10/147,603
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 09/539,333
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA
```


ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 31..1107
OTHER INFORMATION: 5'regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 1108..1289
OTHER INFORMATION: exon A g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 14877..14920
OTHER INFORMATION: exon B g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 18778..18862
OTHER INFORMATION: exon Bbis g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 25593..25740
OTHER INFORMATION: exon C g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29388..29502
OTHER INFORMATION: exon D g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29967..30282
OTHER INFORMATION: exon E g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 64666..64812
OTHER INFORMATION: exon F g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 65505..65853
OTHER INFORMATION: exon G g35018 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 65854..67854
OTHER INFORMATION: 3'regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 94124..94964
OTHER INFORMATION: exon g35017
FEATURE:
NAME/KEY: exon
LOCATION: 201188..201234
OTHER INFORMATION: exon S g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 213818..215818
OTHER INFORMATION: 3'regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene

FEATURE:
NAME/KEY: exon
LOCATION: 216661..216952
OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 234174..234321
OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon MS2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:

```

; NAME/KEY: exon
; LOCATION: 240528..241685
; OTHER INFORMATION: exon M1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240800..240993
; OTHER INFORMATION: exon MS1 complement g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 241686..243685
; OTHER INFORMATION: 5'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 290652..292652
; OTHER INFORMATION: 3'regulatory region g34665 gene

Query Match      11.0%; Score 28.6; DB 7; Length 319608;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy   35  GGGTGAAGCCATGCTCTGAAGGGGATGACGTCCCTTGCTCATCCAAAAACCGTCT 94
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   21445 GCCTGTTTCATTGATGAATAAATTAACAGTAATGCCAGTCTTCACAATGCTGCT 21504

Qy   95  CGGGTGGTGAGAGTCTCGCTGTGTGGGAAGCAGTCA 133
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   21505 CAGGTGTGACTGAAGTTGTGATTGATTGGAAAGTACTCA 21543

RESULT 4
US-11-052-544-28
; Sequence 28, Application US/11052544
; Publication No. US20050255504A1
; GENERAL INFORMATION:
; APPLICANT: PARL, Fritz F.
; TITLE OF INVENTION: METHOD OF DETECTING AN INCREASED
; FILE REFERENCE: 22000.0127U2
; CURRENT APPLICATION NUMBER: US/11/052,544
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: 60/543,866
; PRIOR FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 38703
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note =
; OTHER INFORMATION: synthetic construct
US-11-052-544-28

Query Match      10.5%; Score 27.4; DB 7; Length 38703;
Best Local Similarity 50.4%; Pred. No. 14;
Matches 67; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy   57  GGATGAGTCCCTCTGGCTCATCCACAAAACCGTCTGGGTGGGTGAGGAGTCTCTGC 116
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   15809 GGATTCCCCACATGGGGGTTCAATCATACAAAGTTTCTGGGTGTAGCTGTGAATACTGCC 15868

Qy   117  TGTGTGGGAAGCAGTCAAGTAATATCCCGTCGTGTGGTGAGCGCTCACGACGATTTTG 176
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   15869 TGTGGGTTCAGTATTGAACATCAATGCTACTGTGTGTCAGTCTTACTATTGCCAGTGTCA 15928

Qy   177  TCCGCTGTGCAGA 189
     ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   15929 TCCGTGGTCTCTGA 15941

RESULT 5
US-11-051-568-26
; Sequence 26, Application US/11051568

```

```
/ REFERENCE/DOCKET NUMBER: STK-001CP6C3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/248-7000
/ TELEFAX: 617/248-7100
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1926 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: MURIDAE
/ TISSUE TYPE: EMBRYO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 93..1289
/ OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/ product= "mOP2-PP"
/ notes= "mOP2 CDNA"
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-11-051-568-26

Query Match 10.5%; Score 27.2; DB 7; Length 1926;
Best Local Similarity 53.8%; Pred. No. 6.1; Mismatches 48; Indels 0; Gaps 0;
Matches 56; Conservative 0;

QY 1 GAGTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 401 GAGCTTCGTCAACATGGTGGAGCGGACCGTACCTGGGCTACGAGGCCACACTGGAA 460

QY 61 GAGCTCCCTTCTGGCTCATCCACAAACCGTCTCGGTTGGGTG 104
DB 461 GGAATTCACCTTGGACCTAACCCAGATCCCTGCTGGGGAGGCTG 504

RESULT 6
US-11-121-086-59
; Sequence 59, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 59
; LENGTH: 162537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-59

Query Match 10.5%; Score 27.2; DB 7; Length 162537;
Best Local Similarity 54.3%; Pred. No. 27; Mismatches 23; Indels 0; Gaps 0;
Matches 41; Conservative 0;

QY 15 TGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACGTCCTCTCTGG 74
DB 51382 TTGAGGACCCCATGCTGATGTGTAACCCCTGCTCTGAGTGGTGGCTTCCCTTCAGC 51441

QY 75 CTCA 78
DB 51442 CCA 51445

RESULT 7
US-11-121-086-63
; Sequence 63, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

Query Match 10.2%; Score 26.6; DB 7; Length 169725;
Best Local Similarity 60.3%; Pred. No. 46; Mismatches 29; Indels 0; Gaps 0;
Matches 44; Conservative 0;

QY 74 GCTCATCCACAAACCGTCTCGGTTGGGTGAGGAGTCTCGTGTGTGGGAGCAGTCA 133
DB 16630 GCTTGGGCACACGCGCTCTGAAGTGTGTGTGGGCCCTGCCAGCGTGGTTAGCAGACA 16689

QY 134 GTATAATTCCCGT 146
DB 16690 GAGCTGTCGCGT 16702

RESULT 8
US-11-121-086-62/c
; Sequence 62, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 62
; LENGTH: 200628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-62

Query Match 10.2%; Score 26.6; DB 7; Length 200628;
Best Local Similarity 60.3%; Pred. No. 46; Mismatches 29; Indels 0; Gaps 0;
Matches 44; Conservative 0;

QY 74 GCTCATCCACAAACCGTCTCGGTTGGGTGAGGAGTCTCGTGTGTGGGAGCAGTCA 133
DB 64410 GCTTGGGCACACGCGCTCTGAAGTGTGTGTGGGCCCTGCCAGCGTGGTTAGCAGACA 64351

QY 134 GTATAATTCCCGT 146
DB 64350 GAGCTGTCGCGT 64338

RESULT 9
US-10-750-185-58438/c
; Sequence 58438, Application US/10750185
; Publication No. US2005026603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
```

```

; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58438
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Bovine 19866880937133
US-10-750-185-58438

Query Match      10.2%; Score 26.4; DB 6; Length 1683;
Best Local Similarity 59.2%; Pred. No.11;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY      110 TCCTGGCGTGTCGGGAAGCAGTCAGTATAAATCCCGTCGTGTGTGGTGAGCGCCTCAGCAC 169
DB      284 TCAGGCGAGAGTCGGGAAGCAGTCGGTTTCAGTTCCTTGGTGCTCGGTCTCAATTGGGA 225

QY      170 GTATTTGTCGCGTGTG 185
DB      224 AAATTCACAGCTGAG 209

RESULT 10
US-10-750-185-32595
; Sequence 32595, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32595
; LENGTH: 2078
; TYPE: DNA
; ORGANISM: Bovine 19866880696449
US-10-750-185-32595

Query Match      10.2%; Score 26.4; DB 6; Length 2078;
Best Local Similarity 59.2%; Pred. No.12;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY      43 CCATGGTCTGAAGGGGATGAGTCCTTCCTGGCTCATCCACAAAACCGTCTCGGGTGGG 102
DB      1956 CCAGGGTCCGAGGGAGGGGCTCTGTGACGTCGACGCACTGCACATGCTCAGGTAGG 2015

QY      103 TGAGGAGTCTGGGTG 118
DB      2016 AGAACAGGTCTGCCCG 2031

RESULT 11
US-11-102-240-109
; Sequence 109, Application US/11102240
; Publication No. US20050260647A1

```


THIS PAGE BLANK (USPTO)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
C 1	34.8	13.4	657	10	BX964816	Reverse s
C 2	34.8	13.4	670	9	AZ962384	2M0231111
C 3	34.8	13.4	818	11	CK247935	Reverse s
C 4	33.8	13.0	1092	11	CNS05POX	AL347370 Tetraodon
C 5	33.6	12.9	890	7	C0738666	CO738666 S1LE04C21
C 6	33.2	12.8	571	1	AL872732	AL872732
C 7	33.2	12.8	945	5	BQ919900	AGENCOURT
C 8	33.2	12.8	1508	2	B8615183	BE615183
C 9	33	12.7	1054	2	B6687472	601281124
C 10	32.8	12.6	886	7	CV242705	602639444
C 11	32.6	12.5	287	1	AV072177	WS02515.B
C 12	32.6	12.5	539	3	BP744242	AV072177
C 13	32.6	12.5	561	9	AQ657755	BP744242
C 14	32.6	12.5	1164	10	AG065034	Sheared D
C 15	32.6	12.5	1527	10	AG118154	AG065034 Pan trogl
C 16	32.4	12.5	949	5	BQ931612	AG118154 Pan trogl
C 17	32.2	12.4	251	2	B8425694	BQ931612 AGENCOURT
C 18	32.2	12.4	470	2	B8007696	B8425694
C 19	32.2	12.4	645	7	CK525759	BB807696
C 20	32	12.3	283	2	BB385246	CK525759 rswfs0_00
C 21	32	12.3	483	3	BM691308	BB385246
C 22	32	12.3	645	11	CR840902	BM691308 UI-E-C11-
C 23	32	12.3	645	11	CR840902	CR840902 GROAAA9DH

```

RESULT 2
AZ962384
LOCUS
DEFINITION
2M023111F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M023111 F, genomic survey sequence.
AZ962384
ACCESSION
AZ962384.1 GI:13933511
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 670)
REFERENCE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D. Weiss R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center
COMMENT
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0231 row: L column: 11
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 670.
FEATURES
Location/Qualifiers
1..670
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M023111"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match 13.4%; Score 34.8; DB 9; Length 670;
Best Local Similarity 54.8%; Pred. No. 11;
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 134 GTATAATTCGGTGTGGTGGAGCCCTCAGCAGTATTTCGCGCTGTCGACGCGT 193
|||||
Db 351 GTGTGACTGCGCATGTGTGTGTGCACCAACCCCTATGCTCTGCTGTGCCAACATGT 410
|||||

RESULT 4
CNS05POX/c
LOCUS
DEFINITION
CNS05POX 1092 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
045G18 of library A from Tetraodon nigroviridis, genomic survey
sequence.
AL347370
1
GI:8241140
GSS; genome survey sequence.
Tetraodon nigroviridis
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
REFERENCE
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,

```

```

QY 194 AGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGCAACCCCGCTTGAATT 253
|||||
Db 411 AAAATGAAGATAGCCCTCAGGTGTTGTACGATGATGAAACACACATGTGGGAAA 470
|||||

QY 254 AAAAAC 259
|||||
Db 471 GATAAC 476

RESULT 3
CR247935
LOCUS
DEFINITION
CR247935 818 bp DNA linear GSS 06-JUL-2004
Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN90111, genomic survey sequence.
CR247935
ACCESSION
CR247935.1 GI:50026789
VERSION
GSS; genome survey sequence; MICR.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 818)
REFERENCE
Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
Rogers, J. and Bradley, A.
Direct Submission
JOURNAL
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
FEATURES
Location/Qualifiers
1..818
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN90111"
/clone_lib="MHPN"
ORIGIN
Query Match 13.4%; Score 34.8; DB 11; Length 818;
Best Local Similarity 54.8%; Pred. No. 12;
Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 134 GTATAATTCGGTGTGGTGGAGCCCTCAGCAGTATTTCGCGCTGTCGACGCGT 193
|||||
Db 374 GTGTGACTGTGCATGTGTGTGTGCACCCACCCCTCTATGCTCTGCTGTGCCAACATGT 315
|||||

QY 194 AGTACCAAGGGCTGCACCCCGGTTTTTGTTCACAGCGGAGGCAACCCCGCTTGAATT 253
|||||
Db 314 AAAATGAAGATAGCCCTCAGGTGTTGTACGATGATGAAACACACATGTGGGAAA 255
|||||

QY 254 AAAAAC 259
|||||
Db 254 GATAAC 249

RESULT 4
CNS05POX/c
LOCUS
DEFINITION
CNS05POX 1092 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence T7 end of clone
045G18 of library A from Tetraodon nigroviridis, genomic survey
sequence.
AL347370
1
GI:8241140
GSS; genome survey sequence.
Tetraodon nigroviridis
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
1
REFERENCE
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,

```


Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,P.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645
2

Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,P.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
10899143
3 (bases 1 to 1092)

Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

Location/Qualifiers
1. .1092
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="045G18"
/clone_lib="A"
/note="Genoscope sequence ID : C0AA045BD09C1
end : T7"

Query Match 13.0%; Score 33.8; DB 11; Length 1092;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 35 GGGTGAAGCATGGTCTGAAGGGGATGAGTCCTCTGGCTCATCCACAAAACCGTCT 94
Db 654 GAGTGCAGCCTTCTTTGTAATGGAGAGGTGAACACACCCAAACCTGCCAAGGCGTTCA 595
QY 95 CGGTGGGTGAGGAGTCTGGCTGTGTGGAGACAGTCAGTATAATTCCTCGTGTGTG 154
Db 594 TCGTTGGGAGAGGAGGAGCGCTCTGTGGTAGCAGACAGATTCAGAGCGCTCTGGCAG 535
QY 155 GTGACGCCCTCAGACGATTTTGTCCGCTGTGCAGCGTAGTACCA 200
Db 534 GNTCAGACAACTGGATTTTATTTTCTTTAGCTAATATTGTTTACA 489

RESULT 5
C0738666 890 bp mRNA linear EST 29-JUL-2004
LOCUS SL1E04c21c04f1 squirrel embryo library 1 Spermophilus lateralis
DEFINITION CDNA clone 21c04 5', mRNA sequence.
ACCESSION C0738666
VERSION C0738666.1 GI:50825936
KEYWORDS EST.
SOURCE Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM Spermophilus lateralis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Scluridae; Xerinae; Marmotini; Spermophilus.
1 (bases 1 to 890)
REFERENCE Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
Rogers,J. and Cossins,A.R.
AUTHORS Microarray analysis of transcriptional changes during hibernation
in the golden mantled ground squirrel, Spermophilus lateralis
Unpublished (2004)
CONTACT: Andrew R. Cossins
COMMENT Laboratory for Environmental Gene Regulation

University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZB
Tel: +44 (0)151-795-4510
Fax: +44 (0)151-795-4431
Email: cossins@liv.ac.uk
Vector has been trimmed from this EST.
Plate: 21 row: c column: 04
Seq primer: pf1c T7 (5'-AATACGACTCACTATAGG-3')
High quality sequence stop: 890.

Location/Qualifiers
1. .890
/organism="Spermophilus lateralis"
/mol_type="mRNA"
/db_xref="taxon:76772"
/clone="21c04"
/sex="male and female"
/tissue_type="embryo"
/dev_stage="embryonic"
/lab_host="E.coli Electromax DH10B"
/clone_lib="squirrel embryo library 1"
/note="Vector: pf1c; Site 1: SalI GTCGAG; Site 2: BamHI
GGATCC; Normalized and subtracted cDNA library_prepared
from embryos"

Query Match 12.9%; Score 33.6; DB 7; Length 890;
Best Local Similarity 52.9%; Pred. No. 28;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 12 CCATGCTGATCAGAACCGTTTCGGGTGAAGCCATCGTCTGAAGGGGATGAGTCCTCCTTC 71
Db 572 CCATGCTGCTTGGCTCTCTATCTGCTGAGGTTTCCCGAGTGAGGATGAGATTTTATA 631
QY 72 TGGCTCATCCACAAAACCGTCTCGGTGGGTGAGAGTCCTGGCTGTCTGGGAACGACT 131
Db 632 AATGTGCTCAAAACAGCTTCCACAAGAGTGAAGAGGCGCTGTCTGTTGTTGTAAGGGG 691
QY 132 CAGTATAATTCCTCGTC 147
Db 692 CAGTCCATTTCTGTC 707

RESULT 6
AL872732/c 571 bp mRNA linear EST 03-DEC-2003
LOCUS AL872732 XCC-egg Xenopus tropicalis cDNA clone TEgg135n13 5', mRNA
DEFINITION sequence.
ACCESSION AL872732
VERSION AL872732.2 GI:38666145
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 571)
REFERENCE Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
AUTHORS Sanger Xenopus tropicalis EST project 2001
TITLE TROPICALIS SEQUENCE ID: TEgg135n13.plkSP6
JOURNAL Unpublished (2003)
COMMENT On Sep 15, 2002 this sequence version replaced gi:22892997.

Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TEgg135n13.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli XL1-blue.

FEATURES

source
1..571
location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEggl35n13"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from egg. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 12.8%; Score 33.2; DB 1; Length 571;
Best Local Similarity 61.6%; Pred. No. 34;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 60 TGACCTCCTTCGTCTATCCACAAAACCGCTCGGTGGGTGAGGAGTCTCGGTGT 119
Db |||||
383 TGTGGCTCCTCGGTCTATCCACATACACTGTAAAGATTGGCTTTGGAGCCACAAGTGT 324
QY 120 GTGGGAAGCAGTCAGTATATTCCTCCG 145
Db |||||
323 TGGGTTTGTGGCAGGATTTTCCAG 298

RESULT 7

BQ919900
LOCUS BQ919900 945 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT 10016818 NCI CGAP Mam2 Mus musculus cDNA clone
IMAGE:648062 5', mRNA sequence.

ACCESSION BQ919900

VERSION BQ919900.1 GI:22334586

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.

1 (bases 1 to 945)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14034 row: a column: 23

High quality sequence stop: 624.

Location/Qualifiers

1..945

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:648062"

/tissue_type="tumor, biopsy sample"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI CGAP Mam2"

/note="Organ: mammary; Vector: pCMV-SPT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 12.8%; Score 33.2; DB 5; Length 945;
Best Local Similarity 54.6%; Pred. No. 38;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 74 GCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCGGTGTGTTGGGAAGCAGTCA 133

Db |||||

598 GCACACCCGCAACACCGGCTTTGATGGGTGTGGAGACTTCAGAAAGAGAGTCCGNA 657

QY 134 GTATAATTCGCTCGTGTGTGTGACGCTCACGACGTATTGTCCGTGTGCAGAGCG 192

Db |||||

658 CAGTCACCTCCCACTCTAAAGCCGAGGCTTTTAGACAATTATCTGTGCTGTAGAGAG 716

RESULT 8

BE615183
LOCUS BE615183 1508 bp mRNA linear EST 24-AUG-2000
DEFINITION 601281124F1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3623261 5',
mRNA sequence.

ACCESSION BE615183

VERSION BE615183.1 GI:9896782

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

1 (bases 1 to 1508)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM299 row: o column: 06

High quality sequence stop: 307.

Location/Qualifiers

1..1508

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:3623261"

/tissue_type="adenocarcinoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_39"

/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dt priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed

by Ling Hong in the laboratory of Gerald M. Rubin

(University of California, Berkeley) using ZAP-CDNA

synthesis kit (Stratagene) and Superscript II RT (Life

Technologies)."

ORIGIN

Query Match 12.8%; Score 33.2; DB 2; Length 1508;
Best Local Similarity 56.4%; Pred. No. 41;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 89 CCGTCTCGGGTGGGTGAGAGTCCCTGGGTGTGTGGGAAGCAGTATATTCCTCGTCG 148

Db |||||

1084 CGCCCCCAAAATTTTTTGTGGTGCCTCCCGGGGGGGGGGGGTATCTCCCCGTG 1143

QY 149 TGTGTGGTGACCCCTCACGACGTATTTGTCCGCTGTGCAGAGCGTAGTAC 198

Db |||||

1144 GGGGGGGGTTTCTCTCCCGTTTGTGGTGTCCCTGTGGGGGGGGTACTAC 1193

RESULT 9	RESULT 10
LOCUS BG687472/c	LOCUS CV242705
DEFINITION BG687472 1054 bp mRNA linear EST 01-MAY-2001	DEFINITION WS02515.B21_M24 PT-MB-N-A-15 Populus trichocarpa cDNA clone
ACCESSION BG687472	ACCESSION WS02515_M24 3', mRNA sequence.
VERSION BG687472.1	VERSION CV242705
KEYWORDS EST.	KEYWORDS CV242705.1 GI:52495680
SOURCE Homo sapiens (human)	SOURCE EST.
ORGANISM Homo sapiens	ORGANISM Populus trichocarpa (Populus balsamifera subsp. trichocarpa)
REFERENCE 1 (bases 1 to 1054)	REFERENCE 1 (bases 1 to 886)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/	AUTHORS Ralph S. Cooper, D. Kolosova, N. Oddy, C. Butterfield, Y. Kirkpatrick, R. Liu, J. Palmquist, D. Stott, J. Barber, S. Yang, G. Babak, R. Brown-John, M. Chand, S. Featherstone, R. Masson, A. Mayo, M. Moran, J. Olson, T. Wong, D. Ritland, C.E. Siddiqui, A. Holt, R. Jones, S. Marra, M. Ellis, B.E. Douglas, C. Ritland, K. and Bohlmann, J.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)	TITLE The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries
JOURNAL Unpublished (1999)	JOURNAL Unpublished (2004)
COMMENT Contact: Robert Strausberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: CLONETECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LCM1616 row: h column: 12 High quality sequence stop: 336.	COMMENT Contact: Joerg Bohlmann Genome BC forest genomics program University of British Columbia Michael Smith Laboratories, 6174 University Boulevard, Rm. 237, Vancouver, British Columbia, Canada, V6T 1Z3 Tel: 1-604-822-0282 Fax: 1-604-822-2114 Email: bohlmann@mel.ubc.ca Plate: WS02515 row: M column: 24 High quality sequence stop: 886.
FEATURES	FEATURES
source	source
1..1054	1..886
/organism="Homo sapiens"	/organism="Populus trichocarpa"
/mol_type="mRNA"	/mol_type="mRNA"
/db_xref="taxon:9606"	/cultivar="wild clone"
/clone="IMAGE:4762427"	/db_xref="taxon:3694"
/tissue_type="mucoepidermoid carcinoma"	/clone="WS02515_M24"
/lab_host="DH10B (T1 phage-resistant)"	/sex="Male"
/clone_lib="NIH MGC 59"	/lab_host="E. coli DH10B T1 phage resistant cells"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site: 1: SfiI (ggcgctggcc); Site 2: SfiI (ggccatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGGCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."	/clone_lib="PT-MB-N-A-15"
	/note="Vector: pBluescript II SK (+) XR; Site 1: EcoRI (5' end of cDNA); Site 2: XhoI (3' end of cDNA); Terminal vegetative buds from 20 year old trees harvested near Corvallis, Oregon on September 19th, 2001. cDNA was prepared from 5 micrograms of mRNA and directionally ligated into the pBluescript II SK (+) XR vector using the pBluescript II XR cDNA Library Construction Kit according to manufacturer's instructions with modifications (Stratagene). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation. Normalization was applied according to published methods (Bonaldi M.F. et al. (1996) Genome Research 6(9):791) in order to reduce the abundance of highly expressed transcripts."
ORIGIN	ORIGIN
Query Match 12.7%; Score 33; DB 2; Length 1054;	Query Match 12.6%; Score 32.8; DB 7; Length 886;
Best Local Similarity 50.3%; Pred. No. 45;	Best Local Similarity 53.0%; Pred. No. 50;
Matches 81; Conservative 0; Mismatches 80; Indels 0; Gaps 0;	Matches 70; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 91 GTCCTGGGTTGGAGGAGTCTGGCTGTGTGGAGGAGTGTAGTATTCCTCGTCG 150	QY 24 AGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAGTACGTCCTTCCTGCTATCC 83
DB 810 GTATGGGAAGTTTGAAGAAGGTTGTGTGCGCCATGGAGTTAGGAGGAGTGTAGTACCTG 751	DB 487 AAAGCGGTGGTGGTGGCGCTGTGTGAAGGGTTCATTGACCATTTAAGGTTCTGTAG 546
QY 151 TGTGGTGAGCGCTTCAGACGATTTTGTCCGCTGTGCAGAGCGTAGTACCAAGGCGTGCAC 210	QY 84 AAAACCGTCTCGGTGGGTGAGGAGTCTCGCTGTGGGAGCAGTCAAGTATATTC 143
DB 750 TTTGGTTGGGCGCGCAGAGCCCTGTTGAGCGTAATGTGTAGTAAAGGAGTGTGG 691	DB 547 AGGAAAGTTATGAGTGGCTGTGTGTGAGTGTAGTAAAGGAGCATAACAATATT 606
QY 211 CCGCGTTTGTTCGAAGCGGAGGCAACCCCGCTTTGAA 251	QY 144 CGTCGTGTGTGG 155
DB 690 GAGGGTGTGTGTGAAGGAGGAGGCCGCACACATGGGTA 650	DB 607 GGTGTGTGTGG 618

```

RESULT 11
AV072177/c
LOCUS AV072177 287 bp mRNA linear EST 24-JUN-1999
DEFINITION clone 2200002K22, mRNA sequence.
ACCESSION AV072177
VERSION AV072177.1 GI:5192005
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 287)
Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T.,
Hata, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M.,
Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Niitsuma, H., Oda, H.,
Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y.,
Sugahara, Y., Suzuki, H., Suzuki, H., Tateo, M., Tomaru, Y.,
Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T.,
Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Teukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1..287
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2200002K22"
/sex="male"
/tissue_type="stomach"
/dev_stage="adult"
/clone_lib="Mus musculus stomach C57BL/6J adult"
ORIGIN
Query Match 12.5%; Score 32.6; DB 1; Length 287;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 80; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 9 CGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGCTGAAGGGGATGAGTCCC 68
Db 166 CAATAATATGGTCCAAAACCAACGATGCTGTGGTCCCTTTTGTGCTGGGGATCAAGGTAC 107
QY 69 TTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCGCTGTGGGAAGC 128
Db 106 TCTTTACACATCTTATATNCCGGTGGAAAGTGGGTTTCATGACAGGTCGTGGAAAAGG 47
QY 129 AGTCAGTATAATTCCTGTCGTGTGGTGTGACGCCCTCACGA 168
Db 46 TATGGTTATGACTCCCTCTTTTGAAGGCTGGTCCCTGAAGA 7
RESULT 12
BP744242
LOCUS BP744242 539 bp mRNA linear EST 07-OCT-2004
DEFINITION clone LP13EST52H08, mRNA sequence.

```

```

ACCESSION BP744242
VERSION BP744242.1 GI:53886462
KEYWORDS EST.
SOURCE Lingulodinium polyedrum
ORGANISM Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Lingulodinium.
REFERENCE 1 (bases 1 to 539)
AUTHORS Tanikawa, N., Akimoto, H., Ogoh, K., Chun, W. and Ohmiya, Y.
TITLE Expressed sequence tag analysis of the dinoflagellate Lingulodinium
JOURNAL polyedrum during dark phase
PUBMED Photochem. Photobiol. 80, 31-35 (2004)
COMMENT 15339226
Contact: Yoshihiro Ohmiya
AIST, PRESTO
Midorioka, Ikeda, Osaka 563-8577, JAPAN
Tel: 81-72-751-7997
Fax: 81-72-751-9628
Email: y-ohmiya@aist.go.jp.
FEATURES
source
1..539
/organism="Lingulodinium polyedrum"
/mol_type="mRNA"
/db_xref="taxon:160621"
/clone="LP13EST52H08"
/clone_lib="Lingulodinium polyedrum cDNA"
ORIGIN
Query Match 12.5%; Score 32.6; DB 3; Length 539;
Best Local Similarity 66.2%; Pred. No. 52;
Matches 47; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 178 CGCTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCACAGCGGAGGCA 237
Db 263 CGGGCGCGCGAGCGCGCAGGACAGGCGCCCGCCAGCTCTGTTCCTCAATGGACGCC 322
QY 238 ACCCCCGCTTG 248
Db 323 ACCCCCGCTTG 333
RESULT 13
AQ657755/c
LOCUS AQ657755 561 bp DNA linear GSS 23-JUN-1999
DEFINITION Sheared DNA-21P13.TF Sheared DNA Trypanosoma brucei genomic clone
ACCESSION AQ657755
VERSION AQ657755.1 GI:5165523
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 561)
AUTHORS El-Sayed, N., Zhao, S., Zhao, H., Gill, S., Suh, E., Malek, J., Fujii, C.,
Gerrard, C., Leech, V., de Jong, P., Ullu, E., Melville, S.,
Donelson, J., Frazer, C. and Adams, M.
TITLE Determination of clone end sequences from Trypanosoma brucei GUTat
JOURNAL 10.1 sheared DNA library
COMMENT Unpublished (1999)
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tldb/mdb/tbdb/.
Seq primer: M13-Forward
Class: shotgun.

```

```

FEATURES             Location/Qualifiers
     source
     1..1164
        /organism="Pan troglodytes"
        /mol_type="genomic DNA"
        /db_xref="taxon:9598"
        /clone="PTB-054E17.R"
        /sex="male"
        /cell_type="lymphoblast"
        /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      12.5%; Score 32.6; DB 10; Length 1164;
Best Local Similarity 50.3%; Pred. No. 53;
Matches 80; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY  90  CGTCTCGGGTGGAGAGTCTCGCTGTGTGGGAAGCAGTCAGTATTAATTCCTCGT 149
      |||||
Db  479  CTTACCCAGGCTCGCGGATCTTCTCTTTGTGCAGCGCTGGGTGGCTCGT 420
      |||||

QY  150  GTGTGTGAGGCTCAGCAGTATTTTCCTGTGTGCAGAGGTAGTACCAAGGGTGA 209
      |||||
Db  419  GTAGGGCGGTGGCTGTGCGGCTCTCGCTCTCTTTGTATTGTCAAGTACGCCCGGAGTA 360
      |||||

QY  210  CCCCGTTTTTGTCCAAAGCGAGGCGCAACCCCGCTTG 248
      |||||
Db  359  TGTCTGTTAGTGGCTTACTGCTCGCTTCGCGGCTTG 321
      |||||

RESULT 14
AG065034/c
LOCUS             1164 bp      DNA      linear      GSS 03-NOV-2001
DEFINITION       Pan troglodytes DNA, clone: PTB-054E17.R, genomic survey sequence.
ACCESSION        AG065034
VERSION          AG065034.1 GI:16616836
KEYWORDS         GSS.
SOURCE           Pan troglodytes (chimpanzee)
ORGANISM         Pan troglodytes

REFERENCE
 1  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
    Totoki,Y., Watanabe,H. and Sakaki,Y.
    BAC end sequences of Library PTB
    Unpublished
 2  (bases 1 to 1164)
    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
    Totoki,Y., Watanabe,H. and Sakaki,Y.
    Direct Submission
    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
    and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
    1-7-22 Suehiro-chou,Tsukuba-shi, Ibaraki, Japan
    (E-mail:chimbases@sc.riken.go.jp; URL:http://hgp.gsc.riken.go.jp/,
    Tel:81-45-503-9111, Fax:81-45-503-9170)
    Clones are derived from the chimpanzee BAC library PTB This BAC end
    was generated during the R&D process and may have higher chance of
    clone tracking errors.
    PRIMERS
    Sequencing: M13Rev
    LIBRARY
    Vector      : pKS145
    R.Site 1   : SacI
    R.Site 2   : SacI

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES             Location/Qualifiers
     source
     1..1527
        /organism="Pan troglodytes"
        /mol_type="genomic DNA"
        /db_xref="taxon:9598"
        /clone="PTB-126D09.R"
        /sex="male"
        /cell_type="lymphoblast"
        /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      12.5%; Score 32.6; DB 10; Length 1164;
Best Local Similarity 50.8%; Pred. No. 61;
Matches 62; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY  89  CCGTCTCGGGTGGAGAGTCTCGCTGTGTGGGAAGCAGTCAGTATTAATTCCTCGT 148
      |||||
Db  160  CCTCTGCGTGGAGCGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 101
      |||||

QY  149  TGTGTGGTACGCCCTCAGCAGTATTTGTTCGGCTGTGCAGAGCGTAGTACCAAGGGTGC 208
      |||||
Db  100  TNCCTNNTTCCGCTTACCNCNTATTTTTCGCGCGCGCGCGCGGCGGCGGCGGCGGCGG 41
      |||||

QY  209  AC 210
      ||
Db  40  GC 39

RESULT 15
AG118154
LOCUS             1527 bp      DNA      linear      GSS 03-NOV-2001
DEFINITION       Pan troglodytes DNA, clone: PTB-126D09.R, genomic survey sequence.
ACCESSION        AG118154
VERSION          AG118154.1 GI:16738673
KEYWORDS         GSS.
SOURCE           Pan troglodytes (chimpanzee)
ORGANISM         Pan troglodytes

REFERENCE
 1  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
    Totoki,Y., Watanabe,H. and Sakaki,Y.
    BAC end sequences of Library PTB
    Unpublished
 2  (bases 1 to 1527)
    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
    Totoki,Y., Watanabe,H. and Sakaki,Y.
    Direct Submission
    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
    and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
    1-7-22 Suehiro-chou,Tsukuba-shi, Ibaraki, Japan
    (E-mail:chimbases@sc.riken.go.jp; URL:http://hgp.gsc.riken.go.jp/,
    Tel:81-45-503-9111, Fax:81-45-503-9170)
    Clones are derived from the chimpanzee BAC library PTB This BAC end
    was generated during the R&D process and may have higher chance of
    clone tracking errors.
    PRIMERS
    Sequencing: M13Rev
    LIBRARY
    Vector      : pKS145
    R.Site 1   : SacI
    R.Site 2   : SacI

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES             Location/Qualifiers
     source
     1..1527
        /organism="Pan troglodytes"
        /mol_type="genomic DNA"
        /db_xref="taxon:9598"
        /clone="PTB-126D09.R"
        /sex="male"
        /cell_type="lymphoblast"
        /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

```

Query Match 12.5%; Score 32.6; DB 10; Length 1527;
 Best Local Similarity 74.5%; Pred. No. 64; Mismatches 14; Indels 0; Gaps 0;
 Matches 41; Conservative 0;

Qy 122 GGGAAAGCAGTCAGTATTAATTCCTCGTGTGTGTGTCACGCGCTTCACGACGTTATTG 176
 Db 1081 GTGATGCAGTCAGTATTAAGACACGTCGCGGTGATGAGGCGCTCTACACATATGTG 1135

Search completed: December 7, 2005, 22:24:30
 Job time : 2440 secs